1

PRELIMINARY CONSIDERATIONS

> Still, social science is possible, and needs a strong empirical component. Even statistical technique may prove useful – from time to time.
>
> (Freedman, 1987, p. 125)

Before delving into the complexities and details of the field of applied social statistics, we first touch on some germane philosophical issues that lay at the heart of where statistics fit in the bigger picture of science. Though this book is primarily about applied statistical modeling, the end goal is to use statistical modeling in the context of scientific exploration and discovery. To have an appreciation for how statistics are used in science, one must first have a sense of some essential foundations so that one can situate where statistics finds itself within the larger frame of scientific investigation.

1.1 THE PHILOSOPHICAL BASES OF KNOWLEDGE: RATIONALISTIC VERSUS EMPIRICIST PURSUITS

All knowledge can be said to be based on fundamental philosophical assumptions, and hence empirical knowledge derived from the sciences is no different. There have, historically, been two means by which knowledge is thought to be attained. The *rationalist* derives knowledge primarily from mental, cognitive pursuits. In this sense, "real objects" are those originating from the mind rather than obtained empirically.

1

FIGURE 1.1   Observing the behavior of a pigeon in a Skinner box (Dtarazona, 1998).

The *empiricist*, on the other hand, derives knowledge from *experience,* that is, "objective" reality. To the empiricist, knowledge is in the form of tangible objects in the "real world."

Ideally, science should possess a healthy blend of both perspectives. On the one hand, science should, of course, be grounded in objective objects. The objects one studies should be independent of the psychical realm. A cup of coffee is a cup of coffee regardless of our belief or theory about the existence of the cup. On the other hand, void of *any* rationalist activity, science becomes the study of objects for which we are not allowed to assign meaning. For example, the behavior of a pigeon in a Skinner box[1] (see Figure 1.1) can be documented as to the number of times it presses on the lever for the reward of a food pellet. That the pigeon presses on the lever is an empirical reality. *Why* the pigeon presses on the level is theoretical speculation, of which there could be many competing possibilities. Observing data is fine, but without theory, we have very little "guidance" to either explain current observations or predict new ones. B.F. Skinner's theory of *operant conditioning*—that the pigeon presses the lever because it is *reinforced* to do so—is a prime example of where a wedding of rationalism and empiricism takes place. The theory attempts to *explain* or *account for* the pigeon's behavior. It is a narrative of why the pigeon does what it does.

Of course, theorizing can go too far, much too far. One must be cautious to not "overtheorize" too extensively without acknowledging the absence of empirical backing. Is there anything wrong with hypothesizing that cloudy days are associated with depressive moods? No, so long as you are prepared to provide evidence that may

---

[1]B.F. Skinner was a psychologist known for his theory of operant conditioning within the behaviorist tradition in psychology. One of Skinner's primary investigatory tools was that of observing and recording the conditions that would lead a rat, pigeon, or other animal to press a lever for a food pellet in a small chamber. This chamber came to be known as the *Skinner box*. For a read of Skinner, see Rutherford (2009) and Fancher and Rutherford (2011).

support or contradict your theory. If no evidence exists, you may still theorize, but you should then admit to your audience the lack of current empirical support for your hypothesis.

As an example of recent "heighted theorizing," recall the missing *Malaysia Airlines Flight 370* where the Boeing 777 aircraft vanished, apparently without a trace, during its flight from Kuala Lumpur to Beijing in March 2014. Media were sometimes criticized for proposing numerous theories as to its disappearance, ranging from the plane being flown into a hidden location to it being hijacked or a result of pilot suicide. One theory even speculated that the plane was swallowed by a black hole! Speculation is fine and theorizing is a necessary scientific as well as human activity, so long as one is up front about existent available evidence to support the theory one is advancing. Indeed, one could assign probabilities to competing theories and revise such probabilities as new data become available. This is precisely what *Bayesian* philosophers and statisticians aim to do. A theory should only be considered credible however when empirical reality and the theory coincide (see Figure 1.2). The fit may not be perfect, and seldom if ever is, but when the rational coincides well with the empirical, credibility of the *idea* is tentatively assured, at least until potentially new evidence debunks it.

We must also ensure that our theories are not too *convenient* of narratives fit to data. If you have ever witnessed a sporting event where the deciding point occurred by the lucky bounce of a puck in hockey or the breezy push of a tennis ball in midair, only to hear post-match commentators laud the winning team or individual as suddenly so much *better* than the losing team, then you know what I mean. We must be careful not to exaggerate how well our given theory fits data simply because a few data points went "our way." George Box once said that all models are wrong, but some are useful. It is equally true that *all models are wrong, but some are just silly*. In any scientific endeavor, guard against *falling in love with your theory* or otherwise exaggerating it far beyond what the data suggest. Otherwise, it is no longer a legitimate theory, but rather simply *your brand* and more a product of subjective bias and "career-building" than anything scientific. Indeed, one reason I believe why economic predictions, for instance, are often looked upon with suspicion, is because economists, like psychologists (and theoretical physicists, for that matter), are far too quick to advance theories
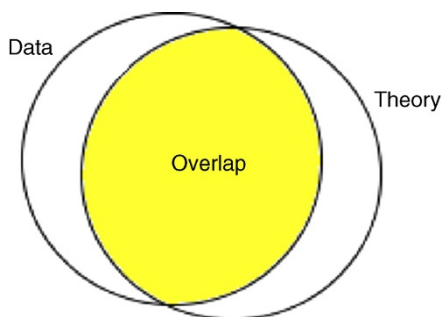
FIGURE 1.2 "Model fit" as an overlap of data with theory.

as though they were facts. "Sexy theories" sound great and may be marketable to uncritical consumers and media, but to good scientists, theories are always only as good as the data that exist to support them. Science is exciting, to be sure, but should not be overly speculative.

1.2   WHAT IS A "MODEL"?

The word "model" is perhaps the most popular word used in textbooks, tutorials, and lectures having anything to do with the application of quantitative methods. Attempting to define just what *is* a model in statistics can be a bit challenging. We discuss the concept by referring to Everitt's definition:

> A description of the assumed structure of a set of observations that can range from a fairly imprecise verbal account to, more usually, a formalized mathematical expression of the process assumed to have generated the observed data. The purpose of such a description is to aid in understanding the data.
>
> (Everitt, 2002, p. 247)

Models are, essentially, and perhaps somewhat crudely, *equations*. They are equations fit to data that attempt to account for how the data came about or were *generated* in the first place. For example, if for every hour a student studied for an exam corresponded to exactly a 1-point increase in a student's grade, the model that would best explain how these data were generated would be a *linear* model. Even if the relationship between hours studied and student grade was not *perfect*, a *perfect* line might still be the "best" summary. Models are often used to account for messy or imperfect data.

Another example of a model is the classic Hebbian version of the *Yerkes–Dodson* curve expressing the relationship between performance and arousal, depicted in Figure 1.3.
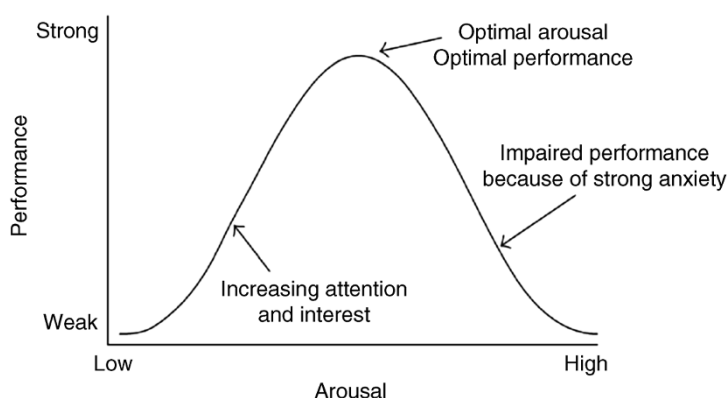

FIGURE 1.3   Hebbian Yerkes–Dodson performance–arousal curve (Diamond et al. 2007).
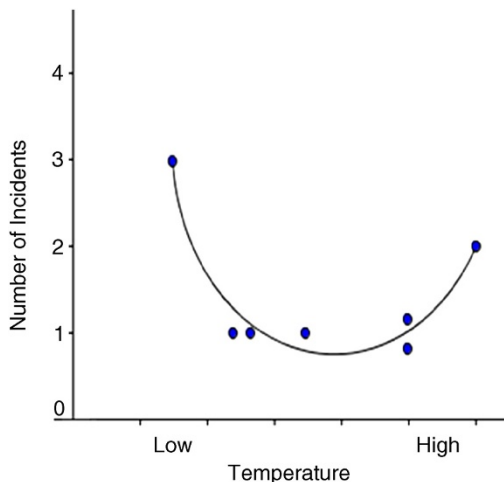
FIGURE 1.4 Number of o-ring incidents on boosters as a function of temperature.

The curve is an inverted "U" shape that provides a useful model relating these two attributes. If one exhibits a very low degree of arousal, performance will be minimal. If one exhibits a very high degree of arousal, performance will likely also suffer. However, if one exhibits a moderate range of arousal, performance will likely be optimal. The model in this case, as in most cases, does not account for *all* the data one might collect. The extent to which it accounts for *most* of the data is the extent to which the model may be, in general, deemed "useful."

As another example of a model, consider the number of o-ring incidents on NASA's space shuttle (the fleet is officially retired now) as a function of temperature (Figure 1.4). At very low or high temperatures, the number of incidents appears to be elevated. A square function seems to adequately model the relationship. Does it account for all points? No. But nonetheless, it provides a fairly good summary of the available data. Some have argued that had NASA had such a model (i.e., the line joining the points) available before *Challenger* was launched on January 28, 1986, the launch may have been delayed and the shuttle and crew saved from disaster.[2] We analyze these data in our chapter on logistic regression.

Why did George Box say that *all models are wrong, some are useful*? The reason is that even if we obtain a perfectly fitting model, there is nothing to say that this is the *only* model that will account for the observed data. Some, such as Fox (1997), even encourage divorcing statistical modeling as accounting for deterministic processes. In discussing the determinants of one's income, for instance, Fox remarks:

> I believe that a statistical model cannot, and is not literally meant to, capture the social
> process by which incomes are "determined" . . . No regression model, not even one

---

[2]See Friendly (2000, pp. 208–211) for an analysis of the o-ring data. See Vaughan (1996) for an account of the social, political, and managerial influences at NASA that were also purportedly responsible for the disaster.

including a residual, can reproduce this process . . . The unfortunate tendency to reify statistical models – to forget that they are descriptive summaries, not literal accounts of social processes – can only serve to discredit quantitative data analysis in the social sciences.

<div align="right">(p. 5)</div>

Indeed, psychological theory, for instance, has advanced numerous models of behavior just as biological theory has advanced numerous theories of human functioning. Two or more competing models may each explain observed data quite well. Sometimes, and unfortunately, the model we adopt may have more to do with our sociological (and even political) preferences than anything to do with whether one is more "correct" than the other. Science is a *human* activity, and often theories that are deemed valid or true have much to do with the *spirit of the times* (the so-called *Zeitgeist*) and what the scientific community will actually accept and tolerate as being *true*.[3]

1.3   SOCIAL SCIENCES VERSUS HARD SCIENCES

There is often a stated distinction between the so-called "soft" sciences and the "hard" sciences (Meehl, 1967). The distinction, as is true in many cases of so many things, is *fuzzy* and *blurry*. The difference between "soft" and "hard" science has usually only to do with the *object* of study, and not with the method of analytical inquiry.

For example, consider what distinguishes the scientist who studies temperature of a human organism from the scientist who studies the self-esteem of adolescents. Their analytical approaches, at their core, will be remarkably similar. They will both measure, collect data, and subject that data to curve-fitting or probabilistic analysis (i.e., statistical modeling). Their *objects*, however, are quite different. Indeed, some may even doubt the *measurability* of something called "self-esteem" in the first place. Is self-esteem *real*? Does it actually exist? At the heart of the distinction, really, is that of *measurement*. Once measurement of an object is agreed upon, the debate between the hard and soft sciences usually vanishes. Both scientists, natural and social, generally aim to do the same thing, and that is to understand and document phenomena and to identify *relations* among phenomena. As Hays (1994) put it, the overreaching goal of science, at its core, is to determine *what goes with what*.

Social science is a courageous attempt. Hard sciences are, in many respects, much easier than the softer social sciences, not necessarily in their subject matter (organic chemistry I have heard *is* difficult), but rather in what they attempt to accomplish. Studying *beats per minute* in an organism is relatively easy. It is not that difficult to measure. Studying something called *intelligence* is much, much harder. Why? Because even arriving at a suitable and agreeable operational definition of what

[3]The reader is strongly encouraged to consult Kuhn's excellent book *The Structure of Scientific Revolutions* in which an eminent philosopher of science argues for what makes some theories more long-standing than others and why some theories drop out of fashion. So-called *paradigm shifts* are present in virtually all sciences. An awareness of such shifts can help one better put "theories of the day" into their proper context.

constitutes intelligence is difficult. Everyone agrees on what "heart rate" means. Fewer people agree on what intelligence really means, even if everyone can agree that some people have more of the mysterious quality than do others. But the study of an object of science should imply that we can actually *measure* it. Intelligence, unlike heart rate, is not easily measured largely because it is a construct open to much scientific criticism and debate. Even if we acknowledge its existence, it is a difficult thing to "tap into."

Given the difficulty in measuring social constructs, should this then mean the social scientist give up and not study the objects of his or her craft? Of course not. But what it does mean is that she must be extremely cautious, conservative, and tentative regarding conclusions drawn from empirical observations. The social scientist must be up front about the weaknesses of her research, and must be very careful not to *overstate* conclusions. For instance, we can measure the extent to which melatonin, a popular sleep aid, reduces the time to sleep onset (i.e., the time it takes to fall asleep). We can perform experimental trials where we give some subjects melatonin and others none, and record who falls to sleep faster. If we keep getting the same results time and time again across a variety of experimental settings, we begin to draw the conclusion that melatonin has a role in decreasing sleep onset. We may not know *why* this is occurring (maybe we do, but I'm pretending for the moment we don't), but we can be reasonably sure the phenomenon exists, that "something" is happening.

Now, contrast the melatonin example with the following question: *Do people of greater intelligence, on average, earn more money than those of lesser intelligence?* We could correlate a measure of intelligence to income, and in this way, we are proceeding in a similar *empirical* (even if not *experimental*, in this case) fashion as would the natural scientist. However, there is a problem. There is a *big* problem. Since only a few consistently agree on what intelligence *is* or how to actually measure it, or even whether it "exists" in the first place, we are unsure of where to even *begin*. Once we agree on what IQ is, how it is measured, and how we will identify and name it, the correlation between IQ and income is as reputable and respectable as the correlation between such variables as height and weight. It is getting to the very *measurement* of IQ that is the initial hard, and skeptics would argue, impossible part. But we know this already from experience. Convincing a parent that her son has an elevated heart rate is much easier than convincing her that her son has a deficit in IQ points. One phenomenon is measurable, the other, perhaps so, but not nearly as easily, or at minimum, *agreeably*.

Our point is that once we agree on the existence, meaning, and measurement of objects, soft science is just as "hard" as the hard sciences. If measurement is not on solid ground, no analytical method of its data will save it. All students of the social sciences should be exposed to in-depth coursework on the theory, philosophy, and importance of *measurement* to their field before advancing to statistical applications on these objects, since it is in the realm of measurement where the true controversies of scientific "reputability" usually lay. For general readable introductions to measurement in psychology and the social sciences, the reader is encouraged to consult Cohen, Swerdlik, and Sturman (2013), Furr and Bacharach (2013), and Raykov and

Marcoulides (2011). For a more deeper and philosophical treatment that includes measurement in the physical sciences as well, refer to Kyburg (2009). McDonald (1999) also provides a relatively technical treatment.

1.4   IS COMPLEXITY A GOOD DEPICTION OF REALITY? ARE MULTIVARIATE METHODS USEFUL?

One of the most prominent advances in social statistics is that of *structural equation modeling* (SEM). With SEM, as we will survey in Chapter 16, one can model complex networks of variables, both *measurable* and *unmeasurable*. Structural equation modeling is indeed one of the most complex statistical methods in the toolkit of the social scientist. However, it is a perfectly fair and reasonable question to ask whether structural equation modeling has helped advance the *cause* of social science. Has it increased our knowledge of social phenomena? Advanced as the tool may be *statistically*, has the tool helped social science build a bigger and better house for itself?

Such questions are open to debate, one that we will not have here. What needs to be acknowledged from the outset however is that statistical complexity has little, if anything, to do with *scientific* complexity or the guarantee of scientific advance. Indeed, the two may even rarely correlate. A classic scenario is that of the graduate student running an independent samples *t*-test on operationally well-defined experimental variables, yet feeling somewhat "guilty" that he used such a "simple" statistical technique. In the laboratory next door, another graduate student is using a complex structural equation model, struggling to make the model identifiable through fixing and freeing parameters at will, yet feeling as though she is more "sophisticated" scientifically as a result of her use of a complex statistical methodology. Hogwash! True, the SEM user may be more sophisticated *statistically* (i.e., SEM *is* harder to understand and implement than *t*-tests), but whether her empirical project is advancing our state of knowledge more than the experimental design of the student using a *t*-test cannot even begin to be evaluated based on the statistical methodology used. It must instead be based on *scientific merit* and the overall strength of the scientific claim. Which scientific contribution is more noteworthy? That is the essential question, not the statistical technique used. The statistics used rarely have anything to do with whether good science versus bad science was performed. Good science is good science, which at times may require statistical analysis as a tool for communicating its findings.

In fact, much of the most rigorous science often requires the most simple and elementary of statistical tools. Students of research can often become dismayed and temporarily disillusioned when they learn that complex statistical methodology, aesthetic and pleasurable on its own that it may be (i.e., SEM models *are* fun to work with), still does not solve their problems. Research-wise, their problems are usually those of *design, controls, and coming up with good experiments, arguments, and ingenious studies*. Their problems are usually not statistical at all, and in this sense, an overemphasis on statistical complexity could actually delay their progress to conjuring up innovative, groundbreaking *scientific* ideas.

The cold hard facts then are that if you have poor design, weak research ideas, and messy measurement of questionable phenomena, your statistical model will provide you with anticlimactic findings, and will be nothing more than an exercise in the old adage "garbage in, garbage out." Quantitative modeling, sophisticated as it has become, has not replaced the need for strict, rigorous experimental controls and good experimental design. Quantitative modeling has not made correlational research somehow more "on par" with the gold standard of experimental studies. Even with the advent of latent variable modeling strategies and methodologies such as confirmatory factor analysis and structural equation modeling, statistics does not purport to "discover" for *real*, hidden variables. Modeling is simply concerned with the partitioning of variability and the estimation of parameters. Beyond that, the remainder of the job of the scientist is to know his or her craft and to design experiments and studies that enlighten and advance our knowledge of a given field. When applied to sound design and thoughtful investigatory practices, statistical modeling *does* partake in this enlightenment, but it does nothing to save the scientist from his or her poorly planned or executed research design. Statistical modeling, complex and enjoyable as it may be on its own, guarantees nothing.

1.5   CAUSALITY

One might say that the ultimate goal of any science is to establish *causal relations*. Nothing suggests a stronger understanding of a scientific field than to be able to speak of causation about the phenomena it studies. However, more difficult than establishing causation in a given research paradigm is that of understanding what causation *means* in the first place. There exist several definitions of causality. Most definitions have at their core that causation is a relation between two events in which the second event is assumed to be a *consequence*, in some sense, of the first event.

For example, if I slip on a banana peel and fall, we might hypothesize that the banana peel *caused* my fall. However, was it the banana peel that caused my fall, or was it the worn-out soles of my shoes that I was wearing that day that caused the fall? Had I been wearing mountain climbers instead of worn-out running shoes, I might not have fallen. Who am I to say the innocent banana peel caused my fall? Causality is *hard*. Even if it seems that *A caused B*, there are usually many variables associated with the problem such that if adjusted or tweaked may threaten the causal claim. Some would say this is simply a trivial philosophical problem of specifying causality and it is "obvious" from the situation that the banana peel caused the fall. Nonetheless, it is clear from such a simple example that causation is in no way an *easy* conclusion to draw. Perhaps this is also why it is extremely difficult to pinpoint true causes of virtually any social behavior. Hindsight is 20/20, but attributing causal attributes with any kind of methodological certainty in violent crimes, for instance, usually turns out to be speculative at best. True, we may accumulate evidence for *prediction*, but equating that with causation is under most circumstances the *wish*, not the reality, of a social theory.

In our brief discussion here, we will not attempt to define causality. Books, dissertations, and treatises have been written exclusively on the topic. At most, what

we can do in the limited space we have is to simply give the following advice to the reader: *If you are going to speak of causation with regard to your research, be prepared to back up your theory of causation to your audience.* It is simply not enough to say *A causes B* without subjecting yourself to at least some of the philosophical issues that accompany such a statement. Otherwise, it is strongly advised that you avoid words like *cause* in hypothesizing or explaining results and findings. *Relations* and *predictions* are much epistemologically "safer" words to use. For a brief but enlightening discussion of causality in the social sciences, see Fox (1997, pp. 3–14). For a more thorough treatment of the subject as it relates to structural equation models, see Mulaik (2009, pp. 63–117).

1.6   THE NATURE OF MATHEMATICS: MATHEMATICS AS A REPRESENTATION OF CONCEPTS

Stewart (1995) said it best when he wrote that the mathematician is not a juggler of numbers, he is a juggler of *concepts*. The greatest ambivalence to learning statistical modeling experienced by students outside (and even inside, I suppose) the mathematical sciences is that of the presumed mathematical complexity involved in such pursuits. Who wants to learn a mathematically-based subject such as statistics when one has *never been good at math*?

The first step in this pursuit is to critically examine assumptions and prior learned beliefs that have become implicit. One way to help "demystify" mathematics and statistics is to challenge your perception of what mathematics and statistics *are* in the first place. It is of great curiosity that so many students claim to dislike mathematics and statistics, yet at the same time cannot verbalize just what mathematics and statistics actually *are*, and then even worse, proceed to engage in real-life activities that utilize very much the same analytical cognitive capacities as would be demanded from doing mathematics and statistics! More than likely, the "dislike" of these subjects has more to do with the *perceptions* one has learned to associate with these subjects than with an inherent ontological disdain for them. Human beings are creatures of psychological *association*. Any dislike of anything without knowing what that thing *is* in the first place is almost akin to disliking a restaurant dish you have never tried. You cannot dislike something until you at least know something about it and open your mind to new possibilities of what *it might be* that you are forming opinions *about*. Not to sound overly "Jamesian," but perhaps you are afraid of mathematics because of your *fear* of it rather than the mathematics itself. If you accept that you are yet unsure of what mathematics is, and will not judge it until you are knowledgeable of it, it may delay derogatory opinion about it. It is only when we assume we know something that we usually feel free to judge and evaluate it. Keep your perceptions open to revision, and what you may find is that what was disliked yesterday curiously becomes likable today, simply because you have now learned more about what that something actually *is*. But to learn more about it, you need to first drop, or at minimum *suspend*, previously held beliefs about it.

The first point is that statistics is not mathematics. Statistics is a discipline in itself that *uses* mathematics, the way physics uses mathematics, and the way that virtually all of the natural and social sciences use mathematics. Mathematics is the tool statisticians use to express their statistical ideas, and statistics is the tool that social scientists use to help make sense of their research findings. The field of theoretical or mathematical statistics is heavily steeped in theorem-building and proofs. Applied statistics, of the kind featured in this book, is definitely not. Thus, any fear of *real mathematics* can be laid to rest, because you will find no such mathematics in this book.

Mathematics and statistics are not "mysterious" things that can only be grasped by those with higher mental faculties. A useful working definition might be that it is a set of well-defined and ever-expanding *rules* based on fundamental assumptions called *axioms*. The axioms of mathematics are typically assumed to be true without needing to be proved. Theorems and other results built on such axioms usually require proof. What is a proof? It is an analytical argument for why a proposition should be considered true. Any given proof usually relies on other theorems that have already been proven to be true. Make no mistake, mathematics is a very deep field of intellectual endeavor and activity. However, expecting something to be deeper than it is can also lead you to just as well not understand it. Sometimes, if you are not understanding something, it may very well be that you are looking far beyond what there is to be understood. If you retreat in your expectations slightly of what there is to see, it sometimes begins to make more sense. Thinking "too deep" where such depth is not required is a peril.

For a general overview of the nature of mathematics, the reader is encouraged to consult Courant, Robbins, and Stewart (1996), and for an excellent introduction to basic mathematical analysis, Labarre (1961). Hamming (1985) is another good introduction to the field of mathematics, as well as Aleksandrov, Kolmogorov, and Lavrent'ev (1999). For more philosophical treatments, the reader should consult Dunham (1994) and Stewart (1995). For an in-depth and very readable history of mathematics, consult Boyer and Merzbach (1991).

1.7   AS A SOCIAL SCIENTIST, HOW MUCH MATHEMATICS DO YOU NEED TO KNOW?

The answer to this question is, of course, *as much as possible*, for working through mathematical problems of any kind can only serve to hone your analytical and deductive abilities. However, that answer is, of course, a naïve if not idealistic one, since there is only so much time available for study and the study of mathematics and statistics must be balanced by your own study of your chosen field.

For example, if the biology student became immersed in mathematics and statistics full-time, then that student would no longer be a student of biology. It can be exceedingly difficult to apply a statistical technique, and interpret the results of such a technique in a field for which you are not familiar. If you are unaware of the substantive objects you are working with, that is, the "stuff" on which the statistics are

being applied, then regardless of your quantitative expertise, you will have difficulty interpreting results. Likewise, if spending too much time computing higher-order derivatives, the student of animal learning, for instance, will have little time remaining to study the learning patterns of the rats he is conditioning, or to speculate on theoretical advancements in his field. Hence, a "happy medium" is required that will balance your study of your substantive area along with the technical quantitative demands of your field of study. Indeed, even for those who specialize exclusively in statistics, the *American Statistical Association* strongly advises aspiring statisticians to choose a field of application. As a researcher, you will be expected to apply modeling techniques that are quite advanced (entire *courses* are devoted to the statistical technique you may be applying), and so you will face the opposite problem, that of choosing to specialize in statistics (to some extent) so that you may better understand the phenomena of your own science. Hence, regardless of whether one is coming from a mathematics or science background, one should aspire for a healthy mix of scientific *and* statistical expertise.

1.8 STATISTICS AND RELATIVITY

Statistical thinking is all about *relativity*. Statistics are not about numbers, they are about *distributions* of numbers (Green, 2000, personal communication). Rarely in statistics, or science for that matter, do we evaluate things in a vacuum.

Consider a very easy example. You board an airplane destined to your favorite vacation spot. How talented is the pilot who is flying your airplane? Is he a "good" pilot or a "bad" pilot? One would hope he is "good enough" to fulfill his duties and ensure your and other passengers' safety. However, when you start thinking like a statistician, you may ponder how good a pilot he is *relative to other pilots*. Where on the curve does your pilot fall? In terms of his or her skill, the pilot of an airplane can be *absolutely* good, but still *relatively* poor. Perhaps that pilot falls on the lower end of the talent curve for pilots. The pilot is still very capable of flying the plane, for he or she has passed an *absolute standard*, but he or she is just not quite as good as most other pilots (see Figure 1.5).

We can come up with a lot of other examples to illustrate the *absolute versus relative* distinction. If someone asked you whether you are intelligent, ego aside, and as statistician, you may respond "relative to who?" Indeed, with a construct like IQ, relativity is all we really have. What does *absolute* intelligence look like? Should our species discover aliens on another planet one day, we may need to revise our definition of intelligence if such are much more (or much less) advanced than we are. Though of course this would assume we have the intelligence to comprehend that their capacities are more than ours, a fact not guaranteed and hence another example of the *trap of relativity*.

Relativity is a benchmark used to evaluate much phenomena, from intelligence to scholastic achievement, to prevalence of depression, and indeed much of human and nonhuman behavior. Understanding that events witnessed could be theorized to have
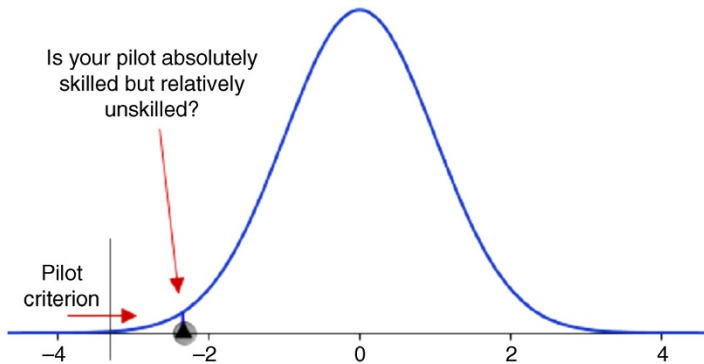
FIGURE 1.5   The "pilot criterion" must be met for *any* pilot to be permitted to fly your plane. However, of those skilled enough to fly, your pilot may still lay at the lower end of the curve. That is, your pilot may be *absolutely* good, but *relatively* poor in terms of skill.

come from known distributions (like the talent distribution of pilots) is a first step to thinking statistically. Most phenomena have distributions, either known or unknown. Statistics, in large part, is a study of such distributions.

1.9   EXPERIMENTAL VERSUS STATISTICAL CONTROL

Perhaps most pervasive in the social science literature is the implicit belief held by many that methods such as regression and analysis of covariance allow one to "control" variables that would otherwise not be controllable in the nonexperimental design. As emphasized throughout this book, statistical methods, whatever the kind, do not provide methods of controlling variables, or "holding variables constant" as it were. Not in the *real* way. To get these kinds of effects, you need a strong and rigorous bullet-proof experimental design.

It is true however that statistical methods do afford a method, in some sense, for presuming (or guessing) *what might have been* had controls been put into place. For instance, if we analyze the correlation between weight and height, it may make sense to hold a factor such as age "constant." That is, we may wish to *partial out* age. However, partialling out the variability due to age in the bivariate correlation is not equivalent to actually *controlling* for age. The truth of the matter is that our statistical control tells us nothing about what would actually *be* had we been able to *truly* control age, or any other factor. As will be elaborated in Chapter 9 on multiple regression, statistical control is not a sufficient "proxy" whatsoever for *experimental* control. Students and researchers must keep this distinction in mind before they throw variables into a statistical model and employ words like "control" (or other *power* and *action* words) when interpreting effects. If you want to truly control variables, to actually hold them *constant*, you will have to do *experiments*.

1.10   STATISTICAL VERSUS PHYSICAL EFFECTS

In the establishment of evidence, either experimental or nonexperimental, it is helpful to consider the distinction between *statistical* versus *physical* effects. To illustrate, consider a medical scientist who wishes to test the hypothesis that the more medication applied to a wound, the faster the wound heals. The statistical question of interest is "*Does the amount of medication predict the rate at which a wound heals?*" A useful statistical model would be a linear regression where amount of medication is the predictor and rate of healing is the response. Of course, one does not *need* a regression analysis to "know" whether something is occurring. The investigator can simply observe whether the wound heals or not, and whether applying more or less medication speeds up or slows down the healing process. The statistical tool in this case is simply used to *model* the relationship, not *determine* whether or not it *exists*. The variable in question is a physical, biological, "real" phenomenon. It exists independent of the statistical model, simply because we can *see* it.

In some areas of social science, however, the very observance of an effect cannot be realized without recourse to the statistics used to model the relationship. For instance, if I correlate self-esteem to intelligence, am I modeling a relationship that I know exists *separate* from the statistical model, or, is the statistical model the only recourse I have to say that the relationship exists in the first place? Because of mediating and moderating relationships in social statistics, an additional variable or two could drastically modify existing coefficients in a model to the point where predictors that had an effect before such inclusion no longer do after. As we will emphasize in our chapters on regression:

> When you change the model, you change parameter estimates, you change effects. You are never, ever, testing individual effects in the model. You are always testing the model, and hence the interpretation of parameter estimates must be within the context of the model.

This is one of the general problems of purely correlational research with nonphysical or "nonorganic" variables. It may be more an exercise in *variance partitioning* than it is in analyzing "true" effects, since the effects in question may be simply statistical artifacts. They may have little other *bases*. Granted, even working with physical or biological variables this can be a problem, but it does not rear its head nearly as much. To reiterate, when we model a physical relationship, we have recourse to that physical relationship independent of the statistical model, because we have evidence that the physical relationship exists independent of the model. If we lost our modeling software, we could still "see" the phenomenon. In many models of social phenomena, however, the addition of one or two covariates in the model can make the relationship of most interest "disappear" and because of the nature of measured variables, we may no longer have physical recourse to justify the original relationship at all, *external to the statistical model*. This is why social models can be very "neurotic," frustrating, and context dependent. Self-esteem may predict achievement in one model, but in another, it does not. Many areas of psychological and political research, for instance, implicitly operate on such grounds. The existence of phenomena is literally "built" on the existence of the

statistical model and often do not exist separate from it, or at least not in an easily observed manner such as the healing of a wound. Social scientists working in such areas, if nothing else, must be aware of this.

1.11 UNDERSTANDING WHAT "APPLIED STATISTICS" MEANS

In the present age of extraordinary computing power, the likes of which will probably seem laughable even after a decade of publication of this book, with a few clicks of the mouse and a software manual, one can obtain a *principal components analysis*, *factor analysis*, *discriminant analysis*, *multiple regression*, and a host of other relatively theoretically advanced statistical techniques in a matter of seconds. The advance of computers and especially easy-to-use software programs has made performing statistical analyses *seemingly* quite easy because even a novice can obtain output from a statistical procedure relatively quickly. One consequence of this, however, is that there seems to have arisen a misunderstanding in some circles that "applied statistics" somehow equates with the idea of "statistics without mathematics" or even worse, "statistics via software."

The word "applied" in *applied statistics* should not be understood to necessarily imply the use of computers. What "applied" should mean is that the focus of the writing is on how to *use* statistics in the context of scientific investigation, often times with demonstrations with real or hypothetical data. Whether that data are analyzed "by hand" or through the use of software does not make one approach more applied than the other. What it does make it is more *computational* compared to the by-hand approach. Indeed, there is a whole field of study known as *computational statistics* that features a variety of software approaches to data analysis. For examples, see Dalgaard (2008), Venables and Ripley (2002), and Friendly (1991, 2000), the latter of these for an emphasis on data visualization. Fox (2002) also provides good coverage of functions in S-Plus and R.

On the opposite end of the spectrum, if a course in statistics is advertised as *not* being applied, then most often it implies that the course is more theoretical or mathematical in nature with a focus on proof and the justification of results. In essence, what this really means is that the course is usually more *abstract* than what would be expected in an applied course. In such theoretical courses, very seldom will one see applications to real data, and instead the course will include proofs of essential statistical theorems and the justification of analytical propositions. Hence, this is the true distinction between *applied* versus *theoretical* courses. The computer has really nothing to do with the distinction other than *facilitating* computation in either field.

REVIEW EXERCISES

1.1. Distinguish between rationalism versus empiricism in accounting for different types of knowledge, and why being a rationalist or empiricist exclusively is usually quite unreasonable and unrealistic.

1.2. Briefly discuss what is meant by a model in scientific research.

1.3. Compare and contrast the social versus "hard" sciences. How are they similar? Different? In this context, discuss the statement "Social science is a courageous attempt."

1.4. Compare and contrast a physical quantity such as weight with a psychological one such as intelligence. How is one more "real" than the other? Can they be considered to be equally real? Why or why not?

1.5. Why would some people say that an attribute such as intelligence is not measurable?

1.6. Discuss George Box's infamous statement "All models are wrong, some are useful."

1.7. Consider an example from your own area of research in which two competing explanations, one simple and one complex, may equally well account for observed data. Then, discuss why the simpler explanation may be preferable to the more complex. Are there instances where the more complex explanation may be preferable to the simpler? Discuss.

1.8. Briefly discuss why using statistical methods to make causal statements about phenomena may be unrealistic and in most cases unattainable. Should the word "cause" be used at all in reference to nonexperimental social research?

1.9. Discuss why it is important to suspend one's beliefs about a subject such as applied statistics or mathematics in order to potentially learn more about it.

1.10. Statistical thinking is about relativity. Discuss what this statement means with reference to the pilot example, then by making up an example of your own.

1.11. Distinguish between experimental versus statistical control, and why understanding the distinction between them is important when interpreting a statistical model.

1.12. Distinguish between statistical versus physical effects and how the effect of a medication treating a wound might be considered different in nature from the correlation between intelligence and self-esteem.

1.13. Distinguish between the domains of applied versus theoretical statistics.

Further Discussion and Activities

1.14 William of Ockham (c. 1287–1347) is known for his infamous principle *Ockham's razor*, which essentially states that all things equal, given competing theories accounting for the same data, the simpler theory is the better theory. In other words, complex explanations for phenomena that could be explained by simpler means are not encouraged. Read Kelly (2007) and evaluate the utility of Ockham's razor as it applies to statistical modeling. Do you agree that the

simpler statistical model is usually preferred over the more complex when it comes to modeling social phenomena? Why or why not?

1.15. Read Kuhn (2012) and discuss what he means by *normal science* and what constitutes *paradigm shifts* in science.

1.16. As briefly discussed in this chapter, *statistical* control is not the same thing as *experimental* control or that of a *control group*. Read Dehue (2005) and provide a brief commentary on what constitutes a *real* control group versus the concept of *statistical* controls.

1.17. In the chapter, potential problems have been briefly discussed regarding using the word *cause* or speaking of *causality* when describing findings in the social and (often) natural sciences. The topic of causality is a philosopher's career and a scientist's methodological nightmare. Like so many other disciplines, epidemiology, the study of diseases in human and other populations, has had to grapple with the issue of causation. For example, if one is to make the statement *smoking causes cancer*, one must be able to defend one's philosophical position in advancing such a claim. Not everyone who smokes gets cancer. Furthermore, some who smoke the most never get the disease, whereas some who smoke the least do. Tobacco companies have historically relied on the fact that *not everyone who smokes gets cancer* as a means for challenging the smoking–cancer "link." As an introduction to these issues, as well as a brief history of causal interpretations, read Morabia (2005). Summarize the historical interpretations of causality, as well as how epidemiology has generally dealt with the problem of causation.

1.18. Models are used across the sciences to help account for empirical observations. How to best relate mathematical models to reality is not at all straightforward. Read Hennig (2009) and discuss his account of the relation between reality and mathematical models. Do you agree with this account? What might be some problems with it?