The Signal and the Noise

The Art and Science of Prediction

NATE SILVER

PENGUIN BOOKS

PENGUIN BOOKS

Published by the Penguin Group
Penguin Books Ltd, 80 Strand, London WC2R 0RL, England
Penguin Group (USA) Inc., 375 Hudson Street, New York, New York 10014, USA
Penguin Group (Canada), 90 Eglinton Avenue East, Suite 700, Toronto, Ontario, Canada M4P 2Y3
(a division of Pearson Penguin Canada Inc.)
Penguin Ireland, 25 St Stephen's Green, Dublin 2, Ireland (a division of Penguin Books Ltd)
Penguin Group (Australia), 707 Collins Street, Melbourne, Victoria 3008, Australia
(a division of Pearson Australia Group Pty Ltd)
Penguin Books India Pvt Ltd, 11 Community Centre, Panchsheel Park, New Delhi – 110 017, India
Penguin Group (NZ), 67 Apollo Drive, Rosedale, Auckland 0632, New Zealand
(a division of Pearson New Zealand Ltd)
Penguin Books (South Africa) (Pty) Ltd, Block D, Rosebank Office Park, 181 Jan Smuts Avenue,
Parktown North, Gauteng 2193, South Africa

Penguin Books Ltd, Registered Offices: 80 Strand, London WC2R 0RL, England

www.penguin.com

First published in the United States of America by The Penguin Press, a member of
Penguin Group (USA) Inc. 2012
First published in Great Britain by Allen Lane 2012
Published in Penguin Books 2013
004

Illustration credits: Figure 4-2: Courtesy of Dr. Tim Parker, University of Oxford; Figure 7-1: From "1918 Influenza: The Mother of All Pandemics" by Jeffery Taubenberger and David Morens, *Emerging Infectious Disease Journal*, vol. 12, no. 1, January 2006, Centers for Disease Control and Prevention; Figures 9-2, 9-3A, 9-3C, 9-4, 9-5, 9-6 and 9-7: By Cburnett, Wikimedia Commons; Figure 12-2: Courtesy of Dr. J. Scott Armstrong, The Wharton School, University of Pennsylvania

Printed in Great Britain by Clays Ltd, St Ives plc

ISBN: 978–0–141–97565–8

www.greenpenguin.co.uk

CONTENTS

INTRODUCTION

This is a book about information, technology, and scientific progress. This is a book about competition, free markets, and the evolution of ideas. This is a book about the things that make us smarter than any computer, and a book about human error. This is a book about how we learn, one step at a time, to come to knowledge of the objective world, and why we sometimes take a step back.

This is a book about prediction, which sits at the intersection of all these things. It is a study of why some predictions succeed and why some fail. My hope is that we might gain a little more insight into planning our futures and become a little less likely to repeat our mistakes.

More Information, More Problems

The original revolution in information technology came not with the microchip, but with the printing press. Johannes Gutenberg's invention in 1440 made information available to the masses, and the explosion of ideas it produced had unintended consequences and unpredictable effects. It was a spark for the Industrial Revolution in 1775,[1] a tipping point in which civilization suddenly went from having made almost no scientific or economic progress for most of its existence to the exponential rates of growth and change that are familiar to us today. It set in motion the events that would produce the European Enlightenment and the founding of the American Republic.

But the printing press would first produce something else: hundreds of years of holy war. As mankind came to believe it could predict its fate and choose its destiny, the bloodiest epoch in human history followed.[2]

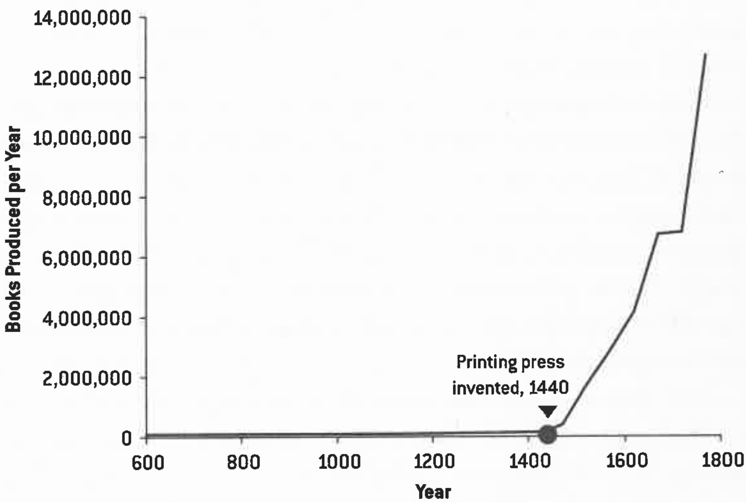
Books had existed prior to Gutenberg, but they were not widely written and they were not widely read. Instead, they were luxury items for the nobility, produced one copy at a time by scribes.[3] The going rate for reproducing a single manuscript was about one florin (a gold coin worth about $200 in today's dollars) per five pages,[4] so a book like the one you're reading now would cost around $20,000. It would probably also come with a litany of transcription errors, since it would be a copy of a copy of a copy, the mistakes having multiplied and mutated through each generation.

This made the accumulation of knowledge extremely difficult. It required heroic effort to prevent the volume of recorded knowledge from actually *decreasing*, since the books might decay faster than they could be reproduced. Various editions of the Bible survived, along with a small number of canonical texts, like from Plato and Aristotle. But an untold amount of wisdom was lost to the ages,[5] and there was little incentive to record more of it to the page.

The pursuit of knowledge seemed inherently futile, if not altogether vain. If today we feel a sense of impermanence because things are changing so rapidly, impermanence was a far more literal concern for the generations before us. There was "nothing new under the sun," as the beautiful Bible verses in Ecclesiastes put it—not so much because everything had been discovered but because everything would be forgotten.[6]

The printing press changed that, and did so permanently and profoundly. Almost overnight, the cost of producing a book decreased by about three hundred times,[7] so a book that might have cost $20,000 in today's dollars instead cost $70. Printing presses spread very rapidly throughout Europe; from Gutenberg's Germany to Rome, Seville, Paris, and Basel by 1470, and then to almost all other major European cities within another ten years.[8] The number of books being produced grew exponentially, increasing by about thirty times in the first century after the printing press was invented.[9] The store of human knowledge had begun to accumulate, and rapidly.

FIGURE I-1: EUROPEAN BOOK PRODUCTION



As was the case during the early days of the World Wide Web, however, the quality of the information was highly varied. While the printing press paid almost immediate dividends in the production of higher quality maps,[10] the bestseller list soon came to be dominated by heretical religious texts and pseudoscientific ones.[11] Errors could now be mass-produced, like in the so-called Wicked Bible, which committed the most unfortunate typo in history to the page: thou *shalt* commit adultery.[12] Meanwhile, exposure to so many new ideas was producing mass confusion. The amount of information was increasing much more rapidly than our understanding of what to do with it, or our ability to differentiate the useful information from the mistruths.[13] Paradoxically, the result of having so much more shared knowledge was increasing isolation along national and religious lines. The instinctual shortcut that we take when we information available to the masses, and the explosion of ideas it produced had unintended consequences and unpredictable effects. It was a spark for the Industrial Revolution in 1775,[1] a tipping point in which civilization suddenly went from having made almost no scientific or economic progress for most of its existence to the exponential rates of growth and change that are familiar to us today. It set in motion the events that would produce the European Enlightenment and the founding of the American Republic.

But the printing press would first produce something else: hundreds of years of holy war. As mankind came to believe it could predict its fate and choose its destiny, the bloodiest epoch in human history followed.[2]

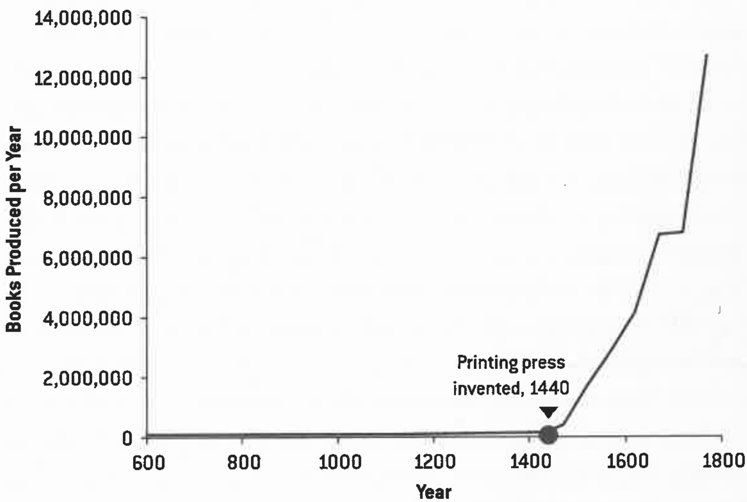
Books had existed prior to Gutenberg, but they were not widely written and they were not widely read. Instead, they were luxury items for the nobility, produced one copy at a time by scribes.[3] The going rate for reproducing a single manuscript was about one florin (a gold coin worth about $200 in today's dollars) per five pages,[4] so a book like the one you're reading now would cost around $20,000. It would probably also come with a litany of transcription errors, since it would be a copy of a copy of a copy, the mistakes having multiplied and mutated through each generation.

This made the accumulation of knowledge extremely difficult. It required heroic effort to prevent the volume of recorded knowledge from actually *decreasing*, since the books might decay faster than they could be reproduced. Various editions of the Bible survived, along with a small number of canonical texts, like from Plato and Aristotle. But an untold amount of wisdom was lost to the ages,[5] and there was little incentive to record more of it to the page.

The pursuit of knowledge seemed inherently futile, if not altogether vain. If today we feel a sense of impermanence because things are changing so rapidly, impermanence was a far more literal concern for the generations before us. There was "nothing new under the sun," as the beautiful Bible verses in Ecclesiastes put it—not so much because everything had been discovered but because everything would be forgotten.[6]

The printing press changed that, and did so permanently and profoundly. Almost overnight, the cost of producing a book decreased by about three hundred times,[7] so a book that might have cost $20,000 in today's dollars instead cost $70. Printing presses spread very rapidly throughout Europe; from Gutenberg's Germany to Rome, Seville, Paris, and Basel by 1470, and then to almost all other major European cities within another ten years.[8] The number of books being produced grew exponentially, increasing by about thirty times in the first century after the printing press was invented.[9] The store of human knowledge had begun to accumulate, and rapidly.

FIGURE I-1: EUROPEAN BOOK PRODUCTION



As was the case during the early days of the World Wide Web, however, the quality of the information was highly varied. While the printing press paid almost immediate dividends in the production of higher quality maps,[10] the bestseller list soon came to be dominated by heretical religious texts and pseudoscientific ones.[11] Errors could now be mass-produced, like in the so-called Wicked Bible, which committed the most unfortunate typo in history to the page: thou *shalt* commit adultery.[12] Meanwhile, exposure to so many new ideas was producing mass confusion. The amount of information was increasing much more rapidly than our understanding of what to do with it, or our ability to differentiate the useful information from the mistruths.[13] Paradoxically, the result of having so much more shared knowledge was increasing isolation along national and religious lines. The instinctual shortcut that we take when we have "too much information" is to engage with it selectively, picking out the parts we like and ignoring the remainder, making allies with those who have made the same choices and enemies of the rest.

The most enthusiastic early customers of the printing press were those who used it to evangelize. Martin Luther's *Ninety-five Theses* were not that radical; similar sentiments had been debated many times over. What was revolutionary, as Elizabeth Eisenstein writes, is that Luther's theses "did not stay tacked to the church door."[14] Instead, they were reproduced at least three hundred thousand times by Gutenberg's printing press[15]—a runaway hit even by modern standards.

The schism that Luther's Protestant Reformation produced soon plunged Europe into war. From 1524 to 1648, there was the German Peasants' War, the Schmalkaldic War, the Eighty Years' War, the Thirty Years' War, the French Wars of Religion, the Irish Confederate Wars, the Scottish Civil War, and the English Civil War—many of them raging simultaneously. This is not to neglect the Spanish Inquisition, which began in 1480, or the War of the Holy League from 1508 to 1516, although those had less to do with the spread of Protestantism. The Thirty Years' War alone killed one-third of Germany's population,[16] and the seventeenth century was possibly the bloodiest ever, with the early twentieth staking the main rival claim.[17]

But somehow in the midst of this, the printing press was starting to produce scientific and literary progress. Galileo was sharing his (censored) ideas, and Shakespeare was producing his plays.

Shakespeare's plays often turn on the idea of fate, as much drama does. What makes them so tragic is the gap between what his characters might like to accomplish and what fate provides to them. The idea of controlling one's fate seemed to have become part of the human consciousness by Shakespeare's time—but not yet the competencies to achieve that end. Instead, those who tested fate usually wound up dead.[18]

These themes are explored most vividly in *The Tragedy of Julius Caesar.* Throughout the first half of the play Caesar receives all sorts of apparent warning signs—what he calls predictions[19] ("beware the ides of March")—that his coronation could turn into a slaughter. Caesar of course ignores these signs, quite proudly insisting that they point to someone else's death—or otherwise reading the evidence selectively. Then Caesar is assassinated.

"[But] men may construe things after their fashion / Clean from the purpose of the things themselves," Shakespeare warns us through the voice of Cicero—good advice for anyone seeking to pluck through their newfound wealth of information. It was hard to tell the signal from the noise. The story the data tells us is often the one we'd like to hear, and we usually make sure that it has a happy ending.

And yet if *The Tragedy of Julius Caesar* turned on an ancient idea of prediction—associating it with fatalism, fortune-telling, and superstition—it also introduced a more modern and altogether more radical idea: that we might interpret these signs so as to gain an advantage from them. "Men at some time are masters of their fates," says Cassius, hoping to persuade Brutus to partake in the conspiracy against Caesar.

The idea of man as master of his fate was gaining currency. The words *predict* and *forecast* are largely used interchangeably today, but in Shakespeare's time, they meant different things. A prediction was what the soothsayer told you; a forecast was something more like Cassius's idea.

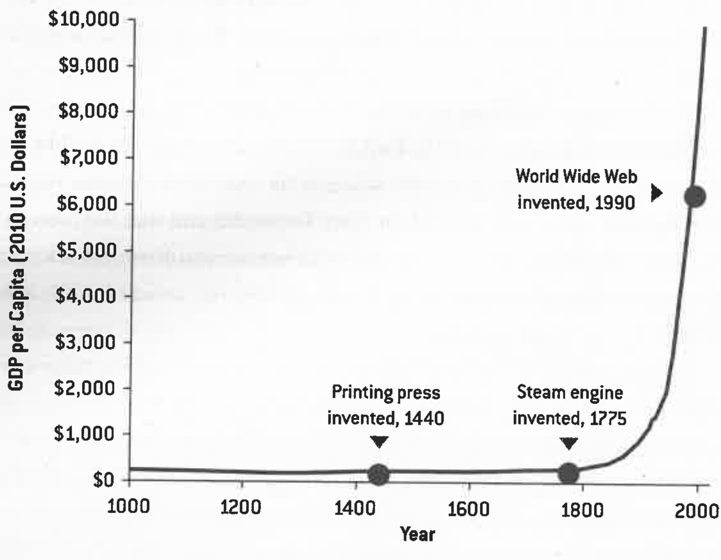
The term *forecast* came from English's Germanic roots,[20] unlike *predict*, which is from Latin.[21] Forecasting reflected the new Protestant worldliness rather than the otherworldliness of the Holy Roman Empire. Making a forecast typically implied planning under conditions of uncertainty. It suggested having prudence, wisdom, and industriousness, more like the way we now use the word *foresight*.[22]

The theological implications of this idea are complicated.[23] But they were less so for those hoping to make a gainful existence in the terrestrial world. These qualities were strongly associated with the Protestant work ethic, which Max Weber saw as bringing about capitalism and the Industrial Revolution.[24] This notion of *forecasting* was very much tied in to the notion of progress. All that information in all those books ought to have helped us to plan our lives and profitably predict the world's course.

The Protestants who ushered in centuries of holy war were learning how to use their accumulated knowledge to change society. The Industrial Revolution largely began in Protestant countries and largely in those with a free press, where both religious and scientific ideas could flow without fear of censorship.[25]

The importance of the Industrial Revolution is hard to overstate. Throughout essentially all of human history, economic growth had proceeded at a rate of perhaps 0.1 percent per year, enough to allow for a very gradual increase in population, but not *any* growth in per capita living standards.[26] And then, suddenly, there was progress when there had been none. Economic growth began to zoom upward much faster than the growth rate of the population, as it has continued to do through to the present day, the occasional global financial meltdown notwithstanding.[27]

FIGURE I-2: GLOBAL PER CAPITA GDP, 1000–2010



The explosion of information produced by the printing press had done us a world of good, it turned out. It had just taken 330 years—and millions dead in battlefields around Europe—for those advantages to take hold.

The Productivity Paradox

We face danger whenever information growth outpaces our understanding of how to process it. The last forty years of human history imply that it can still take a long time to translate information into useful knowledge, and that if we are not careful, we may take a step back in the meantime.

The term "information age" is not particularly new. It started to come into more widespread use in the late 1970s. The related term "computer age" was used earlier still, starting in about 1970.[28] It was at around this time that computers began to be used more commonly in laboratories and academic settings, even if they had not yet become common as home appliances. This time it did not take three hundred years before the growth in information technology began to produce tangible benefits to human society. But it did take fifteen to twenty.

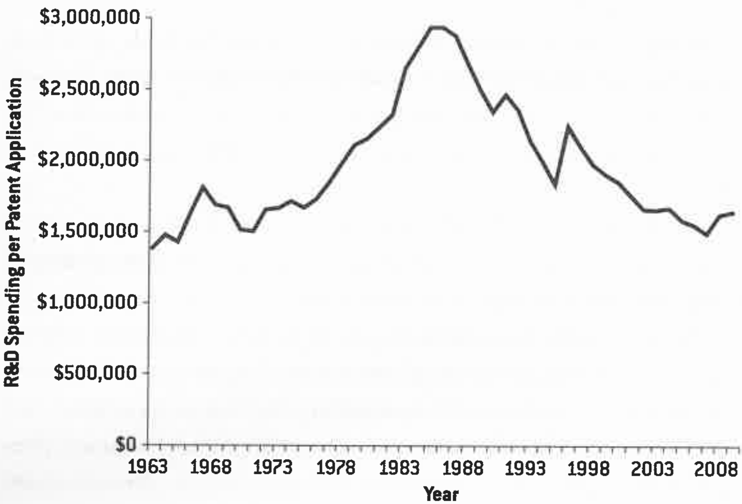
The 1970s were the high point for "vast amounts of theory applied to extremely small amounts of data," as Paul Krugman put it to me. We had begun to use computers to produce models of the world, but it took us some time to recognize how crude and assumption laden they were, and that the precision that computers were capable of was no substitute for predictive accuracy. In fields ranging from economics to epidemiology, this was an era in which bold predictions were made, and equally often failed. In 1971, for instance, it was claimed that we would be able to predict earthquakes within a decade,[29] a problem that we are no closer to solving forty years later.

Instead, the computer boom of the 1970s and 1980s produced a temporary *decline* in economic and scientific productivity. Economists termed this the productivity paradox. "You can see the computer age everywhere but in the productivity statistics," wrote the economist Robert Solow in 1987.[30] The United States experienced four distinct recessions between 1969 and 1982.[31] The late 1980s were a stronger period for our economy, but less so for countries elsewhere in the world.

Scientific progress is harder to measure than economic progress.[32] But one mark of it is the number of patents produced, especially relative to the investment in research and development. If it has become cheaper to produce a new invention, this suggests that we are using our information wisely and are forging it into knowledge. If it is becoming more expensive, this suggests that we are seeing signals in the noise and wasting our time on false leads.

In the 1960s the United States spent about $1.5 million (adjusted for inflation[33]) per patent application[34] by an American inventor. That figure *rose* rather than fell at the dawn of the information age, however, doubling to a peak of about $3 million in 1986.[35]

FIGURE I-3: RESEARCH AND DEVELOPMENT EXPENDITURES PER PATENT APPLICATION



As we came to more realistic views of what that new technology could accomplish for us, our research productivity began to improve again in the 1990s. We wandered up fewer blind alleys; computers began to improve our everyday lives and help our economy. Stories of prediction are often those of long-term progress but short-term regress. Many things that seem predictable over the long run foil our best-laid plans in the meanwhile.

The Promise and Pitfalls of "Big Data"

The fashionable term now is "Big Data." IBM estimates that we are generating 2.5 quintillion bytes of data each day, more than 90 percent of which was created in the last two years.[36]

This exponential growth in information is sometimes seen as a cure-all, as computers were in the 1970s. Chris Anderson, the editor of *Wired* magazine, wrote in 2008 that the sheer volume of data would obviate the need for theory, and even the scientific method.[37]

This is an emphatically pro-science and pro-technology book, and I think of it as a very optimistic one. But it argues that these views are badly mistaken. The numbers have no way of speaking for themselves. We speak for them. We imbue them with meaning. Like Caesar, we may construe them in self-serving ways that are detached from their objective reality.

Data-driven predictions can succeed—and they can fail. It is when we deny our role in the process that the odds of failure rise. Before we demand more of our data, we need to demand more of ourselves.

This attitude might seem surprising if you know my background. I have a reputation for working with data and statistics and using them to make successful predictions. In 2003, bored at a consulting job, I designed a system called PECOTA, which sought to predict the statistics of Major League Baseball players. It contained a number of innovations—its forecasts were probabilistic, for instance, outlining a range of possible outcomes for each player—and we found that it outperformed competing systems when we compared their results. In 2008, I founded the Web site FiveThirtyEight, which sought to forecast the upcoming election. The FiveThirtyEight forecasts correctly predicted the winner of the presidential contest in forty-nine of fifty states as well as the winner of all thirty-five U.S. Senate races.

After the election, I was approached by a number of publishers who wanted to capitalize on the success of books such as *Moneyball* and *Freakonomics* that told the story of nerds conquering the world. This book was conceived of along those lines—as an investigation of data-driven predictions in fields ranging from baseball to finance to national security.

But in speaking with well more than one hundred experts in more than a dozen fields over the course of four years, reading hundreds of journal articles and books, and traveling everywhere from Las Vegas to Copenhagen in pursuit of my investigation, I came to realize that prediction in the era of Big Data was not going very well. I had been lucky on a few levels: first, in having achieved success despite having made many of the mistakes that I will describe, and second, in having chosen my battles well.

Baseball, for instance, is an exceptional case. It happens to be an especially rich and revealing exception, and the book considers why this is so—why a decade after *Moneyball,* stat geeks and scouts are now working in harmony.

The book offers some other hopeful examples. Weather forecasting, which also involves a melding of human judgment and computer power, is one of them. Meteorologists have a bad reputation, but they have made remarkable progress, being able to forecast the landfall position of a hurricane three times more accurately than they were a quarter century ago. Meanwhile, I met poker players and sports bettors who really were beating Las Vegas, and the computer programmers who built IBM's Deep Blue and took down a world chess champion.

But these cases of progress in forecasting must be weighed against a series of failures.

If there is one thing that defines Americans—one thing that makes us exceptional—it is our belief in Cassius's idea that we are in control of our own fates. Our country was founded at the dawn of the Industrial Revolution by religious rebels who had seen that the free flow of ideas had helped to spread not just their religious beliefs, but also those of science and commerce. Most of our strengths and weaknesses as a nation—our ingenuity and our industriousness, our arrogance and our impatience—stem from our unshakable belief in the idea that we choose our own course.

But the new millennium got off to a terrible start for Americans. We had not seen the September 11 attacks coming. The problem was not want of information. As had been the case in the Pearl Harbor attacks six decades earlier, all the signals were there. But we had not put them together. Lacking a proper theory for how terrorists might behave, we were blind to the data and the attacks were an "unknown unknown" to us.

There also were the widespread failures of prediction that accompanied the recent global financial crisis. Our naïve trust in models, and our failure to realize how fragile they were to our choice of assumptions, yielded disastrous results. On a more routine basis, meanwhile, I discovered that we are unable to predict recessions more than a few months in advance, and not for lack of trying. While there has been considerable progress made in controlling inflation, our economic policy makers are otherwise flying blind.

The forecasting models published by political scientists in advance of the 2000 presidential election predicted a landslide 11-point victory for Al Gore.[38] George W. Bush won instead. Rather than being an anomalous result, failures like these have been fairly common in political prediction. A long-term study by Philip E. Tetlock of the University of Pennsylvania found that when political scientists claimed that a political outcome had absolutely *no* chance of occurring, it nevertheless happened about 15 percent of the time. (The political scientists are probably better than television pundits, however.)

There has recently been, as in the 1970s, a revival of attempts to predict earthquakes, most of them using highly mathematical and data-driven techniques. But these predictions envisaged earthquakes that never happened and failed to prepare us for those that did. The Fukushima nuclear reactor had been designed to handle a magnitude 8.6 earthquake, in part because some seismologists concluded that anything larger was impossible. Then came Japan's horrible magnitude 9.1 earthquake in March 2011.

There are entire disciplines in which predictions have been failing, often at great cost to society. Consider something like biomedical research. In 2005, an Athens-raised medical researcher named John P. Ioannidis published a controversial paper titled "Why Most Published Research Findings Are False."[39] The paper studied positive findings documented in peer-reviewed journals: descriptions of successful predictions of medical hypotheses carried out in laboratory experiments. It concluded that most of these findings were likely to fail when applied in the real world. Bayer Laboratories recently confirmed Ioannidis's hypothesis. They could not replicate about *two-thirds* of the positive findings claimed in medical journals when they attempted the experiments themselves.[40]

Big Data *will* produce progress—eventually. How quickly it does, and whether we regress in the meantime, will depend on us.

Why the Future Shocks Us

Biologically, we are not very different from our ancestors. But some stone-age strengths have become information-age weaknesses.

Human beings do not have very many natural defenses. We are not all that fast, and we are not all that strong. We do not have claws or fangs or body armor. We cannot spit venom. We cannot camouflage ourselves. And we cannot fly. Instead, we survive by means of our wits. Our minds are quick. We are wired to detect patterns and respond to opportunities and threats without much hesitation.

"This need of finding patterns, humans have this more than other animals," I was told by Tomaso Poggio, an MIT neuroscientist who studies how our brains process information. "Recognizing objects in difficult situations means generalizing. A newborn baby can recognize the basic pattern of a face. It has been learned by evolution, not by the individual."

The problem, Poggio says, is that these evolutionary instincts sometimes lead us to see patterns when there are none there. "People have been doing that all the time," Poggio said. "Finding patterns in random noise."

The human brain is quite remarkable; it can store perhaps three terabytes of information.[41] And yet that is only about one one-millionth of the information that IBM says is now produced in the world *each day*. So we have to be terribly selective about the information we choose to remember.

Alvin Toffler, writing in the book *Future Shock* in 1970, predicted some of the consequences of what he called "information overload." He thought our defense mechanism would be to simplify the world in ways that confirmed our biases, even as the world itself was growing more diverse and more complex.[42]

Our biological instincts are not always very well adapted to the information-rich modern world. Unless we work *actively* to become aware of the biases we introduce, the returns to additional information may be minimal—or diminishing.

The information overload after the birth of the printing press produced greater sectarianism. Now those different religious ideas could be testified to with more information, more conviction, more "proof"—and less tolerance for dissenting opinion. The same phenomenon seems to be occurring today. Political partisanship began to increase very rapidly in the United States beginning at about the time that Toffler wrote *Future Shock* and it may be accelerating even faster with the advent of the Internet.[43]

These partisan beliefs can upset the equation in which more information will bring us closer to the truth. A recent study in *Nature* found that the *more* informed that strong political partisans were about global warming, the *less* they agreed with one another.[44]

Meanwhile, if the quantity of information is increasing by 2.5 quintillion bytes per day, the amount of *useful* information almost certainly isn't. Most of it is just noise, and the noise is increasing faster than the signal. There are so many hypotheses to test, so many data sets to mine—but a relatively constant amount of objective truth.

The printing press changed the way in which we made mistakes. Routine errors of transcription became less common. But when there was a mistake, it would be reproduced many times over, as in the case of the Wicked Bible.

Complex systems like the World Wide Web have this property. They may not fail as often as simpler ones, but when they fail they fail badly. Capitalism and the Internet, both of which are incredibly efficient at propagating information, create the potential for bad ideas as well as good ones to spread. The bad ideas may produce disproportionate effects. In advance of the financial crisis, the system was so highly levered that a single lax assumption in the credit ratings agencies' models played a huge role in bringing down the whole global financial system.

Regulation is one approach to solving these problems. But I am suspicious that it is an excuse to avoid looking within ourselves for answers. We need to stop, and admit it: we have a prediction problem. We love to predict things—and we aren't very good at it.

The Prediction Solution

If prediction is the central problem of this book, it is also its solution.

Prediction is indispensable to our lives. Every time we choose a route to work, decide whether to go on a second date, or set money aside for a rainy day, we are making a forecast about how the future will proceed—and how our plans will affect the odds for a favorable outcome.

Not all of these day-to-day problems require strenuous thought; we can budget only so much time to each decision. Nevertheless, you are making predictions many times every day, whether or not you realize it.

For this reason, this book views prediction as a shared enterprise rather than as a function that a select group of experts or practitioners perform. It is amusing to poke fun at the experts when their predictions fail. However, we should be careful with our Schadenfreude. To say our predictions are no worse than the experts' is to damn ourselves with some awfully faint praise.

Prediction does play a particularly important role in science, however. Some of you may be uncomfortable with a premise that I have been hinting at and will now state explicitly: we can *never* make perfectly objective predictions. They will *always* be tainted by our subjective point of view.

But this book is emphatically against the nihilistic viewpoint that there is no objective truth. It asserts, rather, that a belief in the objective truth—and a commitment to pursuing it—is the first prerequisite of making better predictions. The forecaster's next commitment is to realize that she perceives it imperfectly.

Prediction is important because it connects subjective and objective reality. Karl Popper, the philosopher of science, recognized this view.[45] For Popper, a hypothesis was not scientific unless it was falsifiable—meaning that it could be tested in the real world by means of a prediction.

What should give us pause is that the few ideas we have tested aren't doing so well, and many of our ideas have not or cannot be tested at all. In economics, it is much easier to test an unemployment rate forecast than a claim about the effectiveness of stimulus spending. In political science, we can test models that are used to predict the outcome of elections, but a theory about how changes to political institutions might affect policy outcomes could take decades to verify.

I do not go as far as Popper in asserting that such theories are therefore unscientific or that they lack any value. However, the fact that the few theories we *can* test have produced quite poor results suggests that many of the ideas we *haven't* tested are very wrong as well. We are undoubtedly living with many delusions that we do not even realize.

But there is a way forward. It is not a solution that relies on half-baked policy ideas—particularly given that I have come to view our political system as a big part of the problem. Rather, the solution requires an attitudinal change.

This attitude is embodied by something called Bayes's theorem, which I introduce in chapter 8. Bayes's theorem is nominally a mathematical formula. But it is really much more than that. It implies that we must think differently about our ideas—and how to test them. We must become more comfortable with probability and uncertainty. We must think more carefully about the assumptions and beliefs that we bring to a problem.

The book divides roughly into halves. The first seven chapters diagnose the prediction problem while the final six explore and apply Bayes's solution.

Each chapter is oriented around a particular subject and describes it in some depth. There is no denying that this is a detailed book—in part because that is often where the devil lies, and in part because my view is that a certain amount of immersion in a topic will provide disproportionately more insight than an executive summary.

The subjects I have chosen are usually those in which there is some publicly shared information. There are fewer examples of forecasters making predictions based on private information (for instance, how a company uses its customer records to forecast demand for a new product). My preference is for topics where you can check out the results for yourself rather than having to take my word for it.

A Short Road Map to the Book

The book weaves between examples from the natural sciences, the social sciences, and from sports and games. It builds from relatively straightforward cases, where the successes and failures of prediction are more easily demarcated, into others that require slightly more finesse.

Chapters 1 through 3 consider the failures of prediction surrounding the recent financial crisis, the successes in baseball, and the realm of political prediction—where some approaches have worked well and others haven't. They should get you thinking about some of the most fundamental questions that underlie the prediction problem. How can we apply our judgment to the data—without succumbing to our biases? When does market competition make forecasts better—and how can it make them worse? How do we reconcile the need to use the past as a guide with our recognition that the future may be different?

Chapters 4 through 7 focus on *dynamic* systems: the behavior of the earth's atmosphere, which brings about the weather; the movement of its tectonic plates, which can cause earthquakes; the complex human interactions that account for the behavior of the American economy; and the spread of infectious diseases. These systems are being studied by some of our best scientists. But dynamic systems make forecasting more difficult, and predictions in these fields have not always gone very well.

Chapters 8 through 10 turn toward solutions—first by introducing you to a sports bettor who applies Bayes's theorem more expertly than many economists or scientists do, and then by considering two other games, chess and poker. Sports and games, because they follow well-defined rules, represent good laboratories for testing our predictive skills. They help us to a better understanding of randomness and uncertainty and provide insight about how we might forge information into knowledge.

Bayes's theorem, however, can also be applied to more existential types of problems. Chapters 11 through 13 consider three of these cases: global warming, terrorism, and bubbles in financial markets. These are hard problems for forecasters and for society. But if we are up to the challenge, we can make our country, our economy, and our planet a little safer.

* * *

The world has come a long way since the days of the printing press. Information is no longer a scarce commodity; we have more of it than we know what to do with. But relatively little of it is useful. We perceive it selectively, subjectively, and without much self-regard for the distortions that this causes. We think we want information when we really want knowledge.

The signal is the truth. The noise is what distracts us from the truth. This is a book about the signal and the noise.

8

LESS AND LESS AND LESS WRONG*

The sports bettor Haralabos "Bob" Voulgaris lives in a gleaming, modernist house in the Hollywood Hills of Los Angeles—all metal and glass, with a pool in the back, like something out of a David Hockney painting. He spends every night from November through June watching the NBA, five games at a time, on five Samsung flat screens (the DirecTV guys had never seen anything like it). He escapes to his condo at Palms Place in Las Vegas whenever he needs a short break, and safaris in Africa when he needs a longer one. In a bad year, Voulgaris makes a million dollars, give or take. In a good year, he might make three or four times that.

So Bob enjoys some trappings of the high life. But he doesn't fit the stereotype of the cigar-chomping gambler in a leisure suit. He does not depend on insider tips, crooked referees, or other sorts of hustles to make his bets. Nor does he have a "system" of any kind. He uses computer simulations, but does not rely upon them exclusively.

What makes him successful is the way that he analyzes information. He is

* The title of this chapter is inspired by a line from the poem "The Road to Wisdom," by the Danish mathematician Piet Hein: "to err and err and err again, but less and less and less."

not just hunting for patterns. Instead, Bob combines his knowledge of statistics with his knowledge of basketball in order to identify meaningful *relationships* in the data.

This requires a lot of hard work—and sometimes a lot of guts. It required a big, calculated gamble to get him to where he is today.

Voulgaris grew up in Winnipeg, Manitoba, a hardworking but frostbitten city located ninety miles north of the Minnesota border. His father had once been quite wealthy—worth about $3 million dollars at his peak—but he blew it all gambling. By the time Voulgaris was twelve, his dad was broke. By the time he was sixteen, he realized that if he was going to get the hell out of Winnipeg, he needed a good education and would have to pay for it himself. So while attending the University of Manitoba, he looked for income wherever he could find it. In the summers, he'd go to the far northern reaches of British Columbia to work as a tree climber; the going rate was seven cents per tree. During the school year, he worked as an airport skycap, shuttling luggage back and forth for Winnipeggers bound for Toronto or Minneapolis or beyond.

Voulgaris eventually saved up to buy out a stake in the skycap company that he worked for and, before long, owned much of the business. By the time he was a college senior, in 1999, he had saved up about $80,000.

But $80,000 still wasn't a *lot* of money, Voulgaris thought—he'd seen his dad win and lose several times that amount many times over. And the job prospects for a philosophy major from the University of Manitoba weren't all that promising. He was looking for a way to accelerate his life when he came across a bet that he couldn't resist.

That year, the Los Angeles Lakers had hired the iconoclastic coach Phil Jackson, who had won six championships with the Chicago Bulls. The Lakers had plenty of talent: their superstar center, the seven-foot-one behemoth Shaquille O'Neal, was at the peak of his abilities, and their twenty-one-year-old guard Kobe Bryant, just four years out of high school, was turning into a superstar in his own right. Two great players—a big man like O'Neal and a scorer like Bryant—has long been a formula for success in the NBA, especially when they are paired with a great coach like Jackson who could manage their outsize egos.

And yet conventional wisdom was skeptical about the Lakers. They had never gotten into a rhythm the previous year, the strike-shortened season of 1998–99, when they churned through three coaches and finished 31-19, eliminated in four straight games by the San Antonio Spurs in the second round of the playoffs. Bryant and O'Neal were in a perpetual feud, with O'Neal apparently jealous that Bryant—still not old enough to drink legally—was on the verge of eclipsing him in popularity, his jersey outselling O'Neal's in Los Angeles sporting goods stores.[1] The Western Conference was strong back then, with cohesive and experienced teams like San Antonio and Portland, and the rap was that the Lakers were too immature to handle them.

When the Lakers were blown out by Portland in the third game of the regular season, with O'Neal losing his cool and getting ejected midway through the game, it seemed to confirm all the worst fears of the pundits and the shock jocks. Even the hometown *Los Angeles Times* rated the Lakers as just the seventh-best team in the NBA[2] and scolded Vegas handicappers for having given them relatively optimistic odds, 4-to-1 against, of winning the NBA title before the season had begun.

Just a couple of weeks into the 1999–2000 regular season, the Vegas bookmakers had begun to buy into the skepticism and had lengthened the Lakers' odds to 6½ to 1, making for a much better payout for anyone who dared to buck the conventional wisdom. Voulgaris was never a big believer in conventional wisdom—it's in large part its shortcomings that make his lifestyle possible—and he thought this was patently insane. The newspaper columnists and the bookies were placing too much emphasis on a small sample of data, ignoring the bigger picture and the context that surrounded it.

The Lakers weren't even playing that badly, Voulgaris thought. They had won five of their first seven games despite playing a tough schedule, adjusting to a new coach, and working around an injury to Bryant, who had hurt his wrist in the preseason and hadn't played yet. The media was focused on their patchy 1998–99 season, which had been interrupted by the strike and the coaching changes, while largely ignoring their 61-21 record under more normal circumstances in 1997–98. Voulgaris had watched a lot of Lakers games: he liked what Jackson was doing with the club. So he placed $80,000—his entire life savings less a little he'd left over for food and tuition—on the Lakers to win the NBA championship. If he won his bet, he'd make half a million dollars. If he lost it, it would be back to working double shifts at the airport.

Initially, Voulgaris's instincts were looking very good. From that point in the season onward, the Lakers won 62 of their remaining 71 contests, including three separate winning streaks of 19, 16, and 11 games. They finished at 67-15, one of the best regular-season records in NBA history. But the playoffs were another matter: the Western Conference was brutally tough in those years, and even with home-court advantage throughout the playoffs—their reward for their outstanding regular season—winning four series in a row would be difficult for the Lakers.

Los Angeles survived a scare against a plucky Sacramento Kings team in the first round of the playoffs, the series going to a decisive fifth game, and then waltzed past Phoenix in the Western Conference Semifinals. But in the next round they drew the Portland Trail Blazers, who had a well-rounded and mature roster led by Michael Jordan's former sidekick—and Jackson's former pupil—Scottie Pippen. Portland would be a rough matchup for the Lakers: although they lacked the Lakers' talent, their plodding, physical style of play often knocked teams out of their rhythm.[3]

The Lakers won the first game of the best-of-seven series fairly easily, but then the roller-coaster ride began. They played inexplicably poorly in the second game in Los Angeles, conceding twenty consecutive points to Portland in the third quarter[4] and losing 106-77, their most lopsided defeat of the season.[5]

The next two games were played at the Rose Garden in Portland, but in Game 3, the Lakers gathered themselves after falling down by as many as thirteen points in the first half, with Bryant swatting away a shot in the final seconds to preserve a two-point victory.[6] They defied gravity again in Game 4, overcoming an eleven-point deficit as O'Neal, a notoriously poor free-throw shooter, made all nine of his attempts.[7] Trailing three games to one in the series, the Trail Blazers were "on death's door," as Jackson somewhat injudiciously put it.[8]

But in the fifth game, at the Staples Center in Los Angeles, the Lakers couldn't shoot the ball straight, making just thirty of their seventy-nine shots in a 96-88 defeat. And in the sixth, back in Portland, they fell out of rhythm early and never caught the tune, as the Blazers marched to a 103-93 win. Suddenly the series was even again, with the deciding Game 7 to be played in Los Angeles.

The prudent thing for a gambler would have been to hedge his bet. For instance, Voulgaris could have put $200,000 on Portland, who were 3-to-2 underdogs, to win Game 7. That would have locked in a profit. If the Blazers won, he would make more than enough from his hedge to cover the loss of his original $80,000 bet, still earning a net profit of $220,000.[9] If the Lakers won instead, his original bet would still pay out—he'd lose his hedge, but net $320,000 from both bets combined.* That would be no half-million-dollar score, but still pretty good.

But there was a slight problem: Voulgaris didn't have $200,000. Nor did he know anybody else who did, at least not anybody he could trust. He was a twenty-three-year-old airport skycap living in his brother's basement in Winnipeg. It was literally Los Angeles or bust.

Early on in the game his chances didn't look good. The Blazers went after O'Neal at every opportunity, figuring they'd either force him to the free-throw line, where every shot was an adventure, or get him into foul trouble instead as he retaliated. Halfway through the second quarter, the strategy was working to a tee, as O'Neal had picked up three fouls and hadn't yet scored from the field. Then Portland went on a ferocious run midway through the third quarter, capped off by a Pippen three-pointer that gave them a sixteen-point lead as boos echoed throughout the Staples Center.[10]

Voulgaris's odds at that point were very long. Rarely did a team[11] that found itself in the Lakers' predicament—down sixteen points with two minutes left to play in the third quarter—come back to win the game; it can be calculated that the odds were about 15-to-1 against their doing so.[12] His bet—his ticket out of Winnipeg—looked all but lost.[13]

But early in the fourth quarter, the downside to Portland's brutally physical style of play suddenly became clear. Their players were beaten-up and fatigued,

---

* This assumes that the Lakers would beat the Indiana Pacers, the Eastern Conference champions, in the NBA Finals, against whom they'd be heavily favored. Voulgaris could have hedged his bet again if he wanted to mitigate that slim risk.

running on fumes and adrenaline. The Lakers were playing before their home crowd, which physiologists have shown provides athletes with an extra burst of testosterone when they need it most.[14] And the Lakers were the younger team, with a more resilient supply of energy.

Portland, suddenly, couldn't hit a shot, going more than six minutes without scoring early in the fourth quarter, right as the Lakers were quickening their pace. L.A. brought their deficit down to single digits, then five points, then three, until Brian Shaw hit a three-pointer to even the score with four minutes left, and Bryant knotted two free-throws a couple of possessions later to give them the lead. Although Portland's shooting improved in the last few minutes, it was too late, as the Lakers made clear with a thunderous alley-oop between their two superstars, Bryant and O'Neal, to clinch the game.

Two weeks later, the Lakers disposed of the Indiana Pacers in efficient fashion to win their first NBA title since the Magic Johnson era. And Bob the skycap was halfway to becoming a millionaire.

How Good Gamblers Think

How did Voulgaris know that his Lakers bet would come through? He didn't. Successful gamblers—and successful forecasters of any kind—do not think of the future in terms of no-lose bets, unimpeachable theories, and infinitely precise measurements. These are the illusions of the sucker, the sirens of his overconfidence. Successful gamblers, instead, think of the future as speckles of probability, flickering upward and downward like a stock market ticker to every new jolt of information. When their estimates of these probabilities diverge by a sufficient margin from the odds on offer, they may place a bet.

The Vegas line on the Lakers at the time that Voulgaris placed his bet, for instance, implied that they had a 13 percent chance of winning the NBA title. Voulgaris did not think the Lakers' chances were 100 percent or even 50 percent—but he was confident they were quite a bit higher than 13 percent. Perhaps more like 25 percent, he thought. If Voulgaris's calculation was right, the bet had a theoretical profit of $70,000.

FIGURE 8-1: HOW VOULGARIS SAW HIS LAKERS BET

Outcome	Probability	Net Profit
Lakers win championship	25%	+$520,000
Lakers do not win championship	75%	–$80,000
Expected profit		+$70,000

If the future exists in shades of probabilistic gray to the forecaster, however, the present arrives in black and white. Bob's theoretical profit of $70,000 consisted of a 25 percent chance of winning $520,000 and a 75 percent chance of losing $80,000 averaged together. Over the long term, the wins and losses will average out: the past and the future, to a good forecaster, can resemble one another more than either does the present since both can be expressed in terms of long-run probabilities. But this was a one-shot bet. Voulgaris needed to have a pretty big edge (the half dozen different reasons he thought the bookies undervalued the Lakers), and a pretty big head on his shoulders, in order to make it.

FIGURE 8-2: THE WORLD THROUGH THE EYES OF A SUCCESSFUL GAMBLER

The Future The Present The Past

Now that Voulgaris has built up a bankroll for himself, he can afford to push smaller edges. He might place three or four bets on a typical night of NBA action. While the bets are enormous by any normal standard they are small compared with his net worth, small enough that he can seem glumly indifferent about them. On the night that I visited, he barely blinked an eye when, on one of the flat screens, the Utah Jazz inserted a seven-foot-two Ukrainian stiff named Kyrylo Fesenko into the lineup, a sure sign that they were conceding the game and that Voulgaris would lose his $30,000 bet on it.

Voulgaris's big secret is that he doesn't have a big secret. Instead, he has a thousand *little* secrets, quanta of information that he puts together one vector at a time. He has a program to simulate the outcome of each game, for instance. But he relies on it only if it suggests he has a very clear edge or it is supplemented by other information. He watches almost every NBA game—some live, some on tape—and develops his own opinions about which teams are playing up to their talent and which aren't. He runs what is essentially his own scouting service, hiring assistants to chart every player's defensive positioning on every play, giving him an advantage that even many NBA teams don't have. He follows the Twitter feeds of dozens of NBA players, scrutinizing every 140-character nugget for relevance: a player who tweets about the club he's going out to later that night might not have his head in the game. He pays a lot of attention to what the coaches say in a press conference and the code that they use: if the coach says he wants his team to "learn the offense" or "play good fundamental basketball," for instance, that might suggest he wants to slow down the pace of the game.

To most people, the sort of things that Voulgaris observes might seem trivial. And in a sense, they are: the big and obvious edges will have been noticed by other gamblers, and will be reflected in the betting line. So he needs to dig a little deeper.

Late in the 2002 season, for instance, Voulgaris noticed that games involving the Cleveland Cavaliers were particularly likely to go "over" the total for the game. (There are two major types of sports bets, one being the point spread and the other being the over-under line or *total*—how many points both teams will score together.) After watching a couple of games closely, he quickly detected the reason: Ricky Davis, the team's point guard and a notoriously selfish player, would be a free agent at the end of the year and was doing everything he could to improve his statistics and make himself a more marketable commodity. This meant running the Cavaliers' offense at a breakneck clip in an effort to create as many opportunities as possible to accumulate points and assists. Whether or not this was good basketball didn't much matter: the Cavaliers were far out of playoff contention.[15] As often as not, the Cavaliers' opponents would be out of contention as well and would be happy to return the favor, engaging them in an unspoken pact to play loose defense and trade baskets in an attempt to improve one another's stats.[16] Games featuring the Cavaliers suddenly went from 192 points per game to 207 in the last three weeks of the season.[17] A bet on the over was not quite a sure thing—there are no sure things—but it was going to be highly profitable.

Patterns like these can sometimes seem obvious in retrospect: *of course* Cavaliers games were going to be higher-scoring if they had nothing left to play for but to improve their offensive statistics. But they can escape bettors who take too narrow-minded a view of the statistics without considering the context that produce them. If a team has a couple of high-scoring games in a row, or even three or four, it usually doesn't mean anything. Indeed, because the NBA has a long season—thirty teams playing eighty-two games each—little streaks like these will occur all the time.[18] Most of them are suckers' bets: they will have occurred for reasons having purely to do with chance. In fact, because the bookmakers will usually have noticed these trends as well, and may have overcompensated for them when setting the line, it will sometimes be smart to bet the other way.

So Voulgaris is *not* just looking for patterns. Finding patterns is easy in any kind of data-rich environment; that's what mediocre gamblers do. The key is in determining whether the patterns represent noise or signal.

But although there isn't any one particular key to why Voulgaris might or might not bet on a given game, there is a particular type of thought process that helps govern his decisions. It is called Bayesian reasoning.

The Improbable Legacy of Thomas Bayes

Thomas Bayes was an English minister who was probably born in 1701—although it may have been 1702. Very little is certain about Bayes's life, even though he lent his name to an entire branch of statistics and perhaps its most famous theorem. It is not even clear that anybody knows what Bayes looked like; the portrait of him that is commonly used in encyclopedia articles may have been misattributed.[19]

What is in relatively little dispute is that Bayes was born into a wealthy family, possibly in the southeastern English county of Hertfordshire. He traveled far away to the University of Edinburgh to go to school, because Bayes was a member of a Nonconformist church rather than the Church of England, and was banned from institutions like Oxford and Cambridge.[20]

Bayes was nevertheless elected as a Fellow of the Royal Society despite a relatively paltry record of publication, where he may have served as a sort of in-house critic or mediator of intellectual debates. One work that most scholars attribute to Bayes—although it was published under the pseudonym John Noon[21]—is a tract entitled "Divine Benevolence."[22] In the essay, Bayes considered the age-old theological question of how there could be suffering and evil in the world if God was truly benevolent. Bayes's answer, in essence, was that we should not mistake our human imperfections for imperfections on the part of God, whose designs for the universe we might not fully understand. "Strange therefore . . . because he only sees the lowest part of this scale, [he] should from hence infer a defeat of happiness in the whole," Bayes wrote in response to another theologian.[23]

Bayes's much more famous work, "An Essay toward Solving a Problem in the Doctrine of Chances,"[24] was not published until after his death, when it was brought to the Royal Society's attention in 1763 by a friend of his named Richard Price. It concerned how we formulate probabilistic beliefs about the world when we encounter new data.

Price, in framing Bayes's essay, gives the example of a person who emerges into the world (perhaps he is Adam, or perhaps he came from Plato's cave) and sees the sun rise for the first time. At first, he does not know whether this is typical or some sort of freak occurrence. However, each day that he survives and the sun rises again, his confidence increases that it is a permanent feature of nature. Gradually, through this purely statistical form of inference, the probability he assigns to his prediction that the sun will rise again tomorrow approaches (although never exactly reaches) 100 percent.

The argument made by Bayes and Price is *not* that the world is intrinsically probabilistic or uncertain. Bayes was a believer in divine perfection; he was also an advocate of Isaac Newton's work, which had seemed to suggest that nature follows regular and predictable laws. It is, rather, a statement—expressed both mathematically and philosophically—about how we learn about the universe: that we learn about it through approximation, getting *closer and closer to the truth* as we gather more evidence.

This contrasted[25] with the more skeptical viewpoint of the Scottish philosopher David Hume, who argued that since we could not be *certain* that the sun would rise again, a prediction that it would was inherently no more rational than one that it wouldn't.[26] The Bayesian viewpoint, instead, regards rationality as a *probabilistic* matter. In essence, Bayes and Price are telling Hume, don't blame nature because you are too daft to understand it: if you step out of your skeptical shell and make some predictions about its behavior, perhaps you will get a little closer to the truth.

Probability and Progress

We might notice how similar this claim is to the one that Bayes made in "Divine Benevolence," in which he argued that we should not confuse our own fallibility for the failures of God. Admitting to our own imperfections is a necessary step on the way to redemption.

However, there is nothing intrinsically religious about Bayes's philosophy.[27] Instead, the most common mathematical expression of what is today recognized as Bayes's theorem was developed by a man who was very likely an atheist,[28] the French mathematician and astronomer Pierre-Simon Laplace.

Laplace, as you may remember from chapter 4, was the poster boy for scientific determinism. He argued that we could predict the universe perfectly—given, of course, that we knew the position of every particle within it and were quick enough to compute their movement. So why is Laplace involved with a theory based on probabilism instead?

The reason has to do with the disconnect between the perfection of nature and our very human imperfections in measuring and understanding it. Laplace was frustrated at the time by astronomical observations that appeared to show anomalies in the orbits of Jupiter and Saturn—they seemed to predict that Jupiter would crash into the sun while Saturn would drift off into outer space.[29] These predictions were, of course, quite wrong, and Laplace devoted much of his life to developing much more accurate measurements of these planets' orbits.[30] The improvements that Laplace made relied on probabilistic inferences[31] in lieu of exacting measurements, since instruments like the telescope were still very crude at the time. Laplace came to view probability as a waypoint between ignorance and knowledge. It seemed obvious to him that a more thorough understanding of probability was essential to scientific progress.[32]

The intimate connection between probability, prediction, and scientific progress was thus well understood by Bayes and Laplace in the eighteenth century—the period when human societies were beginning to take the explosion of information that had become available with the invention of the printing press several centuries earlier, and finally translate it into sustained scientific, technological, and economic progress. The connection is essential—equally to predicting the orbits of the planets and the winner of the Lakers' game. As we will see, science may have stumbled later when a different statistical paradigm, which deemphasized the role of prediction and tried to recast uncertainty as resulting from the errors of our measurements rather than the imperfections in our judgments, came to dominate in the twentieth century.

The Simple Mathematics of Bayes's Theorem

If the philosophical underpinnings of Bayes's theorem are surprisingly rich, its mathematics are stunningly simple. In its most basic form, it is just an algebraic expression with three known variables and one unknown one. But this simple formula can lead to vast predictive insights.

Bayes's theorem is concerned with conditional probability. That is, it tells us the probability that a theory or hypothesis is true *if* some event has happened.

Suppose you are living with a partner and come home from a business trip to discover a strange pair of underwear in your dresser drawer. You will probably ask yourself: what is the probability that your partner is cheating on you? The *condition* is that you have found the underwear; the *hypothesis* you are interested in evaluating is the probability that you are being cheated on. Bayes's theorem, believe it or not, can give you an answer to this sort of question—provided that you know (or are willing to estimate) three quantities:

- First, you need to estimate the probability of the underwear's appearing *as a condition of the hypothesis being true*—that is, you are being cheated upon. Let's assume for the sake of this problem that you are a woman and your partner is a man, and the underwear in question is a pair of panties. If he's cheating on you, it's certainly easy enough to imagine how the panties got there. Then again, even (and perhaps especially) if he is cheating on you, you might expect him to be more careful. Let's say that the probability of the panties' appearing, conditional on his cheating on you, is 50 percent.
- Second, you need to estimate the probability of the underwear's appearing *conditional on the hypothesis being false*. If he isn't cheating, are there some innocent explanations for how they got there? Sure, although not all of them are pleasant (they could be *his* panties). It could be that his luggage got mixed up. It could be that a platonic female friend of his, whom you trust, stayed over one night. The panties could be a gift to you that he forgot to wrap up. None of these theories is inherently untenable, although some verge on dog-ate-my-homework excuses. Collectively you put their probability at 5 percent.
- Third and most important, you need what Bayesians call a *prior probability* (or simply a *prior*). What is the probability you would have assigned to him cheating on you *before* you found the underwear? Of course, it might be hard to be entirely objective about this now that the panties have made themselves known. (Ideally, you establish your priors before you start to examine the evidence.) But sometimes, it is possible to estimate a number like this empirically. Studies have found, for instance, that about 4 percent of married partners cheat on their spouses in any given year,[33] so we'll set that as our prior.

If we've estimated these values, Bayes's theorem can then be applied to establish a *posterior possibility*. This is the number that we're interested in: how likely is it that we're being cheated on, given that we've found the underwear? The calculation (and the simple algebraic expression that yields it) is in figure 8-3.

FIGURE 8-3: BAYES'S THEOREM—UNDERWEAR EXAMPLE

PRIOR PROBABILITY		
Initial estimate of how likely it is that he is cheating on you.	x	4%
A NEW EVENT OCCURS: MYSTERIOUS UNDERWEAR ARE FOUND		
Probability of underwear appearing conditional on his cheating on you.	y	50%
Probability of underwear appearing if he is *not* cheating on you.	z	5%
POSTERIOR PROBABILITY		
Revised estimate of how likely it is that he is cheating on you, given that you've found the underwear.	$\frac{xy}{xy + z(1-x)}$	29%

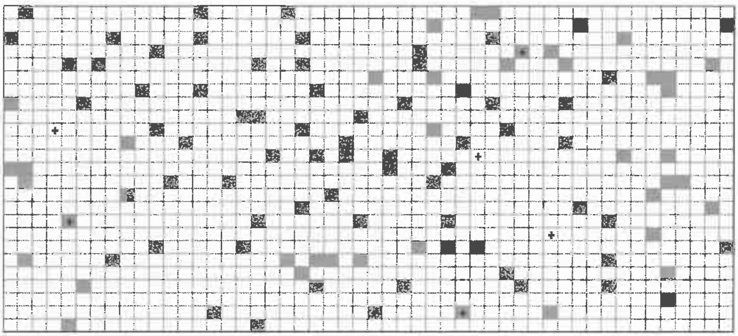
As it turns out, this probability is still fairly low: 29 percent. This may still seem counterintuitive—aren't those panties pretty incriminating? But it stems mostly from the fact that you had assigned a low prior probability to him cheating. Although an innocent man has fewer plausible explanations for the appearance of the panties than a guilty one, you had started out thinking he was an innocent man, so that weighs heavily into the equation.

When our priors are strong, they can be surprisingly resilient in the face of new evidence. One classic example of this is the presence of breast cancer among women in their forties. The chance that a woman will develop breast cancer in her forties is fortunately quite low—about 1.4 percent.[34] But what is the probability if she has a positive mammogram?

Studies show that if a woman does *not* have cancer, a mammogram will incorrectly claim that she does only about 10 percent of the time.[35] If she does have cancer, on the other hand, they will detect it about 75 percent of the time.[36] When you see those statistics, a positive mammogram seems like very bad news indeed. But if you apply Bayes's theorem to these numbers, you'll come to a different conclusion: the chance that a woman in her forties has breast cancer *given that she's had a positive mammogram* is still only about 10 percent. These false positives dominate the equation because very few young women have breast cancer to begin with. For this reason, many doctors recommend that women do not begin getting regular mammograms until they are in their fifties and the prior probability of having breast cancer is higher.[37]

Problems like these are no doubt challenging. A recent study that polled the statistical literacy of Americans presented this breast cancer example to them—and found that just 3 percent of them came up with the right probability estimate.[38] Sometimes, slowing down to look at the problem visually (as in figure 8-4) can provide a reality check against our inaccurate approximations. The visualization makes it easier to see the bigger picture—because breast cancer is so rare in young women, the fact of a positive mammogram is not all that telling.

FIGURE 8-4: BAYES'S THEOREM—MAMMOGRAM EXAMPLE



Women with Breast Cancer (14 of 1000)
- Positive mammogram (true positive) (11 of 14)
- Negative mammorgram (false negative) (3 of 14)

Women Without Breast Cancer (986 of 1000)
- Positive mammogram (false positive) (99 of 986)
- Negative mammorgram (true negative) (887 of 986)

Usually, however, we focus on the newest or most immediately available information, and the bigger picture gets lost. Smart gamblers like Bob Voulgaris have learned to take advantage of this flaw in our thinking. He made a profitable bet on the Lakers in part because the bookmakers placed much too much emphasis on the Lakers' first several games, lengthening their odds of winning the title from 4 to 1 to 6½ to 1, even though their performance was about what you might expect from a good team that had one of its star players injured. Bayes's theorem requires us to think through these problems more carefully and can be very useful for detecting when our gut-level approximations are much too crude.

This is not to suggest that our priors always dominate the new evidence, however, or that Bayes's theorem inherently produces counterintuitive results. Sometimes, the new evidence is so powerful that it overwhelms everything else, and we can go from assigning a near-zero probability of something to a near-certainty of it almost instantly.

Consider a somber example: the September 11 attacks. Most of us would have assigned almost no probability to terrorists crashing planes into buildings in Manhattan when we woke up that morning. But we recognized that a terror attack was an obvious possibility once the first plane hit the World Trade Center. And we had no doubt we were being attacked once the second tower was hit. Bayes's theorem can replicate this result.

For instance, say that before the first plane hit, our estimate of the possibility of a terror attack on tall buildings in Manhattan was just 1 chance in 20,000, or 0.005 percent. However, we would also have assigned a very low probability to a plane hitting the World Trade Center by accident. This figure can actually be estimated empirically: in the previous 25,000 days of aviation over Manhattan[39] prior to September 11, there had been two such accidents: one involving the Empire State Building in 1945 and another at 40 Wall Street in 1946. That would make the possibility of such an accident about 1 chance in 12,500 on any given day. If you use Bayes's theorem to run these numbers (figure 8-5a), the probability we'd assign to a terror attack increased from 0.005 percent to 38 percent the moment that the first plane hit.

FIGURE 8-5A: BAYES'S THEOREM—TERROR ATTACK EXAMPLE

PRIOR PROBABILITY		
Initial estimate of how likely it is that terrorists would crash planes into Manhattan skyscrapers.	x	0.005%
A NEW EVENT OCCURS: FIRST PLANE HITS WORLD TRADE CENTER		
Probability of plane hitting if terrorists are attacking Manhattan skyscrapers.	y	100%
Probability of plane hitting if terrorists are *not* attacking Manhattan skyscrapers (i.e. an accident).	z	0.008%
POSTERIOR PROBABILITY		
Revised estimate of probability of terror attack, given first plane hitting World Trade Center.	$\frac{xy}{xy + z(1-x)}$	38%

The idea behind Bayes's theorem, however, is not that we update our probability estimates just once. Instead, we do so continuously as new evidence presents itself to us. Thus, our posterior probability of a terror attack after the first plane hit, 38 percent, becomes our *prior* possibility before the second one did. And if you go through the calculation again, to reflect the second plane hitting the World Trade Center, the probability that we were under attack becomes a near-certainty—99.99 percent. One accident on a bright sunny day in New York was unlikely enough, but a second one was almost a literal impossibility, as we all horribly deduced.

FIGURE 8-5B: BAYES'S THEOREM—TERROR ATTACK EXAMPLE

PRIOR PROBABILITY		
Revised estimate of probability of terror attack, given first plane hitting World Trade Center.	x	38%
A NEW EVENT OCCURS: SECOND PLANE HITS WORLD TRADE CENTER		
Probability of plane hitting if terrorists are attacking Manhattan skyscrapers.	y	100%
Probability of plane hitting if terrorists are *not* attacking Manhattan skyscrapers (i.e. an accident).	z	0.008%
POSTERIOR PROBABILITY		
Revised estimate of probability of terror attack, given second plane hitting World Trade Center.	$\frac{xy}{xy + z(1-x)}$	99.99%

I have deliberately picked some challenging examples—terror attacks, cancer, being cheated on—because I want to demonstrate the breadth of problems to which Bayesian reasoning can be applied. Bayes's theorem is not any kind of magic formula—in the simple form that we have used here, it consists of nothing more than addition, subtraction, multiplication, and division. We have to provide it with information, particularly our estimates of the prior probabilities, for it to yield useful results.

However, Bayes's theorem does require us to think probabilistically about the world, even when it comes to issues that we don't like to think of as being matters of chance. This does not require us to have taken the position that the world is intrinsically, *metaphysically* uncertain—Laplace thought everything from the orbits of the planets to the behavior of the smallest molecules was governed by orderly Newtonian rules, and yet he was instrumental in the development of Bayes's theorem. Rather, Bayes's theorem deals with *epistemological* uncertainty—the limits of our knowledge.

The Problem of False Positives

When we fail to think like Bayesians, false positives are a problem not just for mammograms but for all of science. In the introduction to this book, I noted the work of the medical researcher John P. A. Ioannidis. In 2005, Ioannidis published an influential paper, "Why Most Published Research Findings Are False,"[40] in which he cited a variety of statistical and theoretical arguments to claim that (as his title implies) the *majority* of hypotheses deemed to be true in journals in medicine and most other academic and scientific professions are, in fact, false.

Ioannidis's hypothesis, as we mentioned, looks to be one of the true ones; Bayer Laboratories found that they could not replicate about *two-thirds* of the positive findings claimed in medical journals when they attempted the experiments themselves.[41] Another way to check the veracity of a research finding is to see whether it makes accurate predictions in the real world—and as we have seen throughout this book, it very often does not. The failure rate for predictions made in entire fields ranging from seismology to political science appears to be extremely high.

"In the last twenty years, with the exponential growth in the availability of information, genomics, and other technologies, we can measure millions and millions of potentially interesting variables," Ioannidis told me. "The expectation is that we can use that information to make predictions work for us. I'm not saying that we haven't made any progress. Taking into account that there are a couple of million papers, it would be a shame if there wasn't. But there are obviously not a couple of million discoveries. Most are not really contributing much to generating knowledge."

This is why our predictions may be *more* prone to failure in the era of Big Data. As there is an exponential increase in the amount of available information, there is likewise an exponential increase in the number of hypotheses to investigate. For instance, the U.S. government now publishes data on about 45,000 economic statistics. If you want to test for relationships between all combinations of two pairs of these statistics—is there a causal relationship between the bank prime loan rate and the unemployment rate in Alabama?—that gives you literally one billion hypotheses to test.*

But the number of *meaningful* relationships in the data—those that speak to causality rather than correlation and testify to how the world really works—is orders of magnitude smaller. Nor is it likely to be increasing at nearly so fast a rate as the information itself; there isn't any more truth in the world than there was before the Internet or the printing press. Most of the data is just noise, as most of the universe is filled with empty space.

Meanwhile, as we know from Bayes's theorem, when the underlying incidence of something in a population is low (breast cancer in young women; truth in the sea of data), false positives can dominate the results if we are not careful. Figure 8-6 represents this graphically. In the figure, 80 percent of true scientific hypotheses are correctly deemed to be true, and about 90 percent of false hypotheses are correctly rejected. And yet, because true findings are so rare, about two-thirds of the findings deemed to be true are actually false!

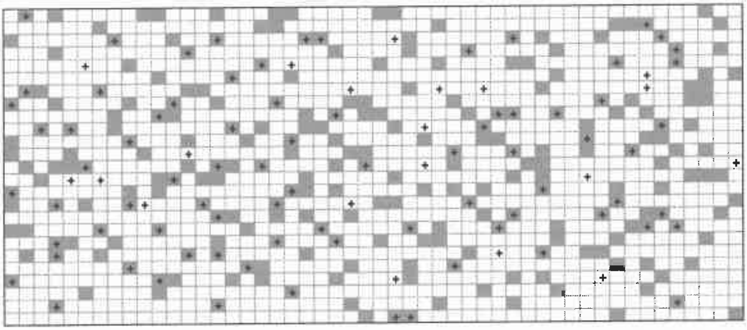
Unfortunately, as Ioannidis figured out, the state of published research in *most* fields that conduct statistical testing is probably very much like what you see in figure 8-6.† Why is the error rate so high? To some extent, this entire book represents an answer to that question. There are many reasons for it—some having to do with our psychological biases, some having to do with common methodological errors, and some having to do with misaligned incentives. Close to the root of the problem, however, is a flawed type of statistical thinking that these researchers are applying.

---

* The number of possible combinations is calculated as 45,000 times 44,999 divided by two, which is 1,012,477,500.

† One difference is that the negative findings are probably kept in a file drawer rather than being published (about 90 percent of the papers published in academic journals today document positive findings rather than negative ones). However, that does not mask the problem of false positives in the findings that do make it to publication.

FIGURE 8-6: A GRAPHICAL REPRESENTATION OF FALSE POSITIVES



True Hypotheses (100 of 1000)
- Positive statistical test (true positive) (80 of 100)
- Negative statistical test (false negative) (20 of 100)

False Hypotheses (900 of 1000)
- Positive statistical test (false positive) (180 of 900)
- Negative statistical test (true negative) (720 of 900)

When Statistics Backtracked from Bayes

Perhaps the chief intellectual rival to Thomas Bayes—although he was born in 1890, almost 120 years after Bayes's death—was an English statistician and biologist named Ronald Aylmer (R. A.) Fisher. Fisher was a much more colorful character than Bayes, almost in the English intellectual tradition of Christopher Hitchens. He was handsome but a slovenly dresser,[42] always smoking his pipe or his cigarettes, constantly picking fights with his real and imagined rivals. He was a mediocre lecturer but an incisive writer with a flair for drama, and an engaging and much-sought-after dinner companion. Fisher's interests were wide-ranging: he was one of the best biologists of his day and one of its better geneticists, but was an unabashed elitist who bemoaned the fact that the poorer classes were having more offspring than the intellectuals.[43] (Fisher dutifully had eight children of his own.)

Fisher is probably more responsible than any other individual for the statistical methods that remain in wide use today. He developed the terminology of the statistical significance test and much of the methodology behind it. He was also no fan of Bayes and Laplace—Fisher was the first person to use the term

"Bayesian" in a published article, and he used it in a derogatory way,[44] at another point asserting that the theory "must be wholly rejected."[45]

Fisher and his contemporaries had no problem with the formula called Bayes's theorem *per se,* which is just a simple mathematical identity. Instead, they were worried about how it might be applied. In particular, they took issue with the notion of the Bayesian prior.[46] It all seemed too subjective: we have to stipulate, in advance, how likely we think something is before embarking on an experiment about it? Doesn't that cut against the notion of objective science?

So Fisher and his contemporaries instead sought to develop a set of statistical methods that they hoped would free us from any possible contamination from bias. This brand of statistics is usually called "frequentism" today, although the term "Fisherian" (as opposed to Bayesian) is sometimes applied to it.[47]

The idea behind frequentism is that uncertainty in a statistical problem results exclusively from collecting data among just a sample of the population rather than the whole population. This makes the most sense in the context of something like a political poll. A survey in California might sample eight hundred people rather than the eight million that will turn out to vote in an upcoming election there, producing what's known as sampling error. The margin of error that you see reported alongside political polls is a measure of this: exactly how much error is introduced because you survey eight hundred people in a population of eight million? The frequentist methods are designed to quantify this.

Even in the context of political polling, however, sampling error does not always tell the whole story. In the brief interval between the Iowa Democratic caucus and New Hampshire Democratic Primary in 2008, about 15,000 people were surveyed[48] in New Hampshire—an enormous number in a small state, enough that the margin of error on the polls was theoretically just plus-or-minus 0.8 percent. The actual error in the polls was about ten times that, however: Hillary Clinton won the state by three points when the polls had her losing to Barack Obama by eight. Sampling error—the *only* type of error that frequentist statistics directly account for—was the least of the problem in the case of the New Hampshire polls.

Likewise, some polling firms consistently show a bias toward one or another party:[49] they could survey all 200 million American adults and they still wouldn't get the numbers right. Bayes had these problems figured out 250 years ago. If you're using a biased instrument, it doesn't matter how many measurements you take—you're aiming at the wrong target.

Essentially, the frequentist approach toward statistics seeks to wash its hands of the reason that predictions most often go wrong: human error. It views uncertainty as something intrinsic to the experiment rather than something intrinsic to our ability to understand the real world. The frequentist method also implies that, as you collect more data, your error will eventually approach zero: this will be both necessary and sufficient to solve any problems. Many of the more problematic areas of prediction in this book come from fields in which useful data is sparse, and it is indeed usually valuable to collect more of it. However, it is hardly a golden road to statistical perfection if you are not using it in a sensible way. As Ioannidis noted, the era of Big Data only seems to be worsening the problems of false positive findings in the research literature.

Nor is the frequentist method particularly objective, either in theory or in practice. Instead, it relies on a whole host of assumptions. It usually presumes that the underlying uncertainty in a measurement follows a bell-curve or normal distribution. This is often a good assumption, but not in the case of something like the variation in the stock market. The frequentist approach requires defining a sample population, something that is straightforward in the case of a political poll but which is largely arbitrary in many other practical applications. What "sample population" was the September 11 attack drawn from?

The bigger problem, however, is that the frequentist methods—in striving for immaculate statistical procedures that can't be contaminated by the researcher's bias—keep him hermetically sealed off from the real world. These methods discourage the researcher from considering the underlying context or plausibility of his hypothesis, something that the Bayesian method demands in the form of a prior probability. Thus, you will see apparently serious papers published on how toads can predict earthquakes,[50] or how big-box stores like Target beget racial hate groups,[51] which apply frequentist tests to produce "statistically significant" (but manifestly ridiculous) findings.

Data Is Useless Without Context

Fisher mellowed out some toward the end of his career, occasionally even praising Bayes.[52] And some of the methods he developed over his long career (although not the ones that are in the widest use today) were really compromises between Bayesian and frequentist approaches. In the last years of his life, however, Fisher made a grievous error of judgment that helps to demonstrate the limitations of his approach.

The issue concerned cigarette smoking and lung cancer. In the 1950s, a large volume of research—some of it using standard statistical methods and some using Bayesian ones[53]—claimed there was a connection between the two, a connection that is of course widely accepted today.

Fisher spent much of his late life fighting against these conclusions, publishing letters in prestigious publications including *The British Medical Journal* and *Nature.*[54] He did not deny that the statistical relationship between cigarettes and lung cancer was fairly strong in these studies, but he claimed it was a case of correlation mistaken for causation, comparing it to a historical correlation between apple imports and marriage rates in England.[55] At one point, he argued that lung cancer caused cigarette smoking and not the other way around[56]—the idea, apparently, was that people might take up smoking for relief from their lung pain.

Many scientific findings that are commonly accepted today would have been dismissed as hooey at one point. This was sometimes because of the cultural taboos of the day (such as in Galileo's claim that the earth revolves around the sun) but at least as often because the data required to analyze the problem did not yet exist. We might let Fisher off the hook if, it turned out, there was not compelling evidence to suggest a linkage between cigarettes and lung cancer by the 1950s. Scholars who have gone back and looked at the evidence that existed at the time have concluded, however, that there was plenty of it—a wide variety of statistical and clinical tests conducted by a wide variety of researchers in a wide variety of contexts demonstrated the causal relationship between them.[57] The idea was quickly becoming the scientific consensus.

So why did Fisher dismiss the theory? One reason may have been that he was a paid consultant of the tobacco companies.[58] Another may have been that he was a lifelong smoker himself. And Fisher liked to be contrarian and controversial, and disliked anything that smacked of puritanism. In short, he was biased, in a variety of ways.

But perhaps the bigger problem is the way that Fisher's statistical philosophy tends to conceive of the world. It emphasizes the objective purity of the experiment—every hypothesis could be tested to a perfect conclusion if only enough data were collected. However, in order to achieve that purity, it denies the need for Bayesian priors or any other sort of messy real-world context. These methods neither require nor encourage us to think about the plausibility of our hypothesis: the idea that cigarettes cause lung cancer competes on a level playing field with the idea that toads predict earthquakes. It is, I suppose, to Fisher's credit that he recognized that correlation does not always imply causation. However, the Fisherian statistical methods do not encourage us to think about *which* correlations imply causations and which ones do not. It is perhaps no surprise that after a lifetime of thinking this way, Fisher lost the ability to tell the difference.

Bob the Bayesian

In the Bayesian worldview, prediction is the yardstick by which we measure progress. We can perhaps never know the truth with 100 percent certainty, but making correct predictions is the way to tell if we're getting closer.

Bayesians hold the gambler in particularly high esteem.[59] Bayes and Laplace, as well as other early probability theorists, very often used examples from games of chance to explicate their work. (Although Bayes probably did not gamble much himself,[60] he traveled in circles in which games like cards and billiards were common and were often played for money.) The gambler makes predictions (good), and he makes predictions that involve estimating probabilities (great), and when he is willing to put his money down on his predictions (even better), he discloses his beliefs about the world to everyone else. The most practical definition of a Bayesian prior might simply be the odds at which you are willing to place a bet.*

And Bob Voulgaris is a particularly Bayesian type of gambler. He likes betting on basketball precisely because it is a way to test himself and the accuracy of his theories. "You could be a general manager in sports and you could be like, Okay, I'll get this player and I'll get that player," he told me toward the end of our interview. "At the end of the day you don't really know if you're right or wrong. But at the end of the day, the end of the season, I know if I'm right or wrong because I know if I'm winning money or I'm losing it. That's a pretty good validation."

Voulgaris soaks up as much basketball information as possible because everything could potentially shift his probability estimates. A professional sports bettor like Voulgaris might place a bet only when he thinks he has at least a 54 percent chance of winning it. This is just enough to cover the "vigorish" (the cut a sportsbook takes on a winning wager), plus the risk associated with putting one's money into play. And for all his skill and hard work—Voulgaris is among the best sports bettors in the world today—he still gets only about 57 percent of his bets right. It is just exceptionally difficult to do much better than that.

A small piece of information that improves Voulgaris's estimate of his odds from 53 percent to 56 percent can therefore make all the difference. This is the sort of narrow margin that gamblers, whether at the poker table or in the stock market, make their living on. Fisher's notion of statistical significance, which uses arbitrary cutoffs devoid of context† to determine what is a "significant" finding and what isn't,[61] is much too clumsy for gambling.

But this is not to suggest that Voulgaris avoids developing *hypotheses* around what he's seeing in the statistics. (The problem with Fisher's notion of

---

* Or more properly, the odds you would set as a betting line so as to be indifferent between either side of the bet. Most Bayesians do require that priors avoid what is called a Dutch book, where the odds are incoherent. If you establish a set of prior probabilities on each of the thirty teams winning the NBA championship, they have to add up to 100 percent exactly since this represents an exhaustive set of possibilities.

† It has been found that because 95 percent confidence in a statistical test is Fisher's traditional dividing line between "signficiant" and "insignificant," researchers are much more likely to report findings that statistical tests classify as 95.1 percent certain than those they classify as 94.9 percent certain—a practice that seems more superstitious than scientific.

hypothesis testing is not with having hypotheses but with the way Fisher recommends that we test them.)[62] In fact, this is critical to what Voulgaris does. Everyone can see the statistical patterns, and they are soon reflected in the betting line. The question is whether they represent signal or noise. Voulgaris forms hypotheses from his basketball knowledge so that he might tell the difference more quickly and more accurately.

Voulgaris's approach to betting basketball is one of the purer distillations of the scientific method that you're likely to find (figure 8-7). He observes the world and asks questions: why are the Cleveland Cavaliers so frequently going over on the total? He then gathers information on the problem, and formulates a hypothesis: the Cavaliers are going over because Ricky Davis is in a contract year and is trying to play at a fast pace to improve his statistics. The difference between what Voulgaris does and what a physicist or biologist might do is that he demarcates his predictions by placing bets on them, whereas a scientist would hope to validate her prediction by conducting an experiment.

FIGURE 8-7: SCIENTIFIC METHOD

Step in Scientific Method[63]	**Sports Betting Example**
Observe a phenomenon	Cavaliers games are frequently going over the game total.
Develop a hypothesis to explain the phenomenon	Cavaliers games are going over because Ricky Davis is playing for a new contract and trying to score as many points as possible.
Formulate a prediction from the hypothesis	Davis's incentives won't change until the end of the season. Therefore: (i) he'll continue to play at a fast pace, and, (ii) future Cavaliers games will continue to be high-scoring as a result.
Test the prediction	Place your bet.

If Voulgaris can develop a strong hypothesis about what he is seeing in the data, it can enable him to make more aggressive bets. Suppose, for instance, that Voulgaris reads some offhand remark from the coach of the Denver Nuggets about wanting to "put on a good show" for the fans. This is probably just idle chatter, but it *might* imply that the team will start to play at a faster pace in order to increase ticket sales. If this hypothesis is right, Voulgaris might expect that an over bet on Nuggets games will win 70 percent of the time as opposed to the customary 50 percent. As a consequence of Bayes's theorem, the stronger Voulgaris's belief in his hypothesis, the more quickly he can begin to make profitable bets on Nuggets games. He might be able to do so after watching just a game or two, observing whether his theory holds in practice—quickly enough that Vegas will have yet to catch on. Conversely, he can avoid being distracted by statistical patterns, like the Lakers' slow start in 1999, that have little underlying meaning but which other handicappers might mistake for a signal.

The Bayesian Path to Less Wrongness

But are Bob's probability estimates subjective or objective? That is a tricky question.

As an empirical matter, we all have beliefs and biases, forged from some combination of our experiences, our values, our knowledge, and perhaps our political or professional agenda. One of the nice characteristics of the Bayesian perspective is that, in explicitly acknowledging that we have prior beliefs that affect how we interpret new evidence, it provides for a very good *description* of how we react to the changes in our world. For instance, if Fisher's prior belief was that there was just a 0.00001 percent chance that cigarettes cause lung cancer, that helps explain why all the evidence to the contrary couldn't convince him otherwise. In fact, there is nothing prohibiting you under Bayes's theorem from holding beliefs that you believe to be *absolutely* true. If you hold there is a 100 percent probability that God exists, or a 0 percent probability, then under Bayes's theorem, *no* amount of evidence could persuade you otherwise.

I'm not here to tell you whether there are things you should believe with *absolute* and *unequivocal* certainty or not.* But perhaps we should be more

---

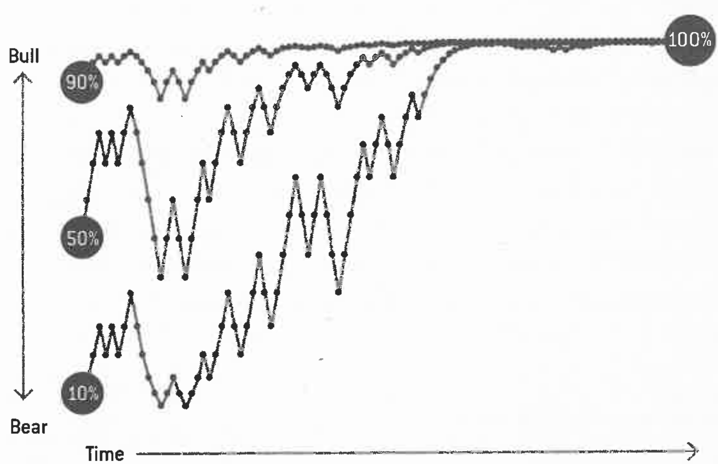
* Although bear in mind that one of the conclusions of this book is that people are overconfident; we probably have too many beliefs that tend toward the 0 percent or 100 percent end of the spectrum.

honest about declaiming these. Absolutely nothing useful is realized when one person who holds that there is a 0 percent probability of something argues against another person who holds that the probability is 100 percent. Many wars—like the sectarian wars in Europe in the early days of the printing press—probably result from something like this premise.

This does not imply that all prior beliefs are equally correct or equally valid. But I'm of the view that we can never achieve perfect objectivity, rationality, or accuracy in our beliefs. Instead, we can strive to be *less* subjective, *less* irrational, and *less* wrong. Making predictions based on our beliefs is the best (and perhaps even the only) way to test ourselves. If objectivity is the concern for a greater truth beyond our personal circumstances, and prediction is the best way to examine how closely aligned our personal perceptions are with that greater truth, the most objective among us are those who make the most accurate predictions. Fisher's statistical method, which saw objectivity as residing within the confines of a laboratory experiment, is less suitable to this task than Bayesian reasoning.

One property of Bayes's theorem, in fact, is that our beliefs should converge toward one another—and toward the truth—as we are presented with more evidence over time. In figure 8-8, I've worked out an example wherein three inves-

FIGURE 8-8: BAYESIAN CONVERGENCE



tors are trying to determine whether they are in a bull market or a bear market. They start out with very different beliefs about this—one of them is optimistic, and believes there's a 90 percent chance of a bull market from the outset, while another one is bearish and says there's just a 10 percent chance. Every time the market goes up, the investors become a little more bullish relative to their prior, while every time it goes down the reverse occurs. However, I set the simulation up such that, although the fluctuations are random on a day-to-day basis, the market increases 60 percent of the time over the long run. Although it is a bumpy road, eventually all the investors correctly determine that they are in a bull market with almost (although not *exactly*, of course) 100 percent certainty.

In theory, science should work this way. The notion of scientific consensus is tricky, but the idea is that the opinion of the scientific community converges *toward* the truth as ideas are debated and new evidence is uncovered. Just as in the stock market, the steps are not always forward or smooth. The scientific community is often too conservative about adapting its paradigms to new evidence,[64] although there have certainly also been times when it was too quick to jump on the bandwagon. Still, provided that everyone is on the Bayesian train,* even incorrect beliefs and quite wrong priors are revised toward the truth in the end.

Right now, for instance, we may be undergoing a paradigm shift in the statistical methods that scientists are using. The critique I have made here about the flaws of Fisher's statistical approach is neither novel nor radical: prominent scholars in fields ranging from clinical psychology[65] to political science[66] to ecology[67] have made similar arguments for years. But so far there has been little fundamental change.

Recently, however, some well-respected statisticians have begun to argue that frequentist statistics should no longer be taught to undergraduates.[68] And some professions have considered banning Fisher's hypothesis test from their journals.[69] In fact, if you read what's been written in the past ten years, it's hard to find anything that *doesn't* advocate a Bayesian approach.

---

* And that they don't hold priors that they believe to be *exactly* 100 percent true or *exactly* 0 percent true; these will not and *cannot* change under Bayes's theorem.

Bob's money is on Bayes, too. He does not literally apply Bayes's theorem every time he makes a prediction. But his practice of testing statistical data in the context of hypotheses and beliefs derived from his basketball knowledge is very Bayesian, as is his comfort with accepting probabilistic answers to his questions.

It will take some time for textbooks and traditions to change. But Bayes's theorem holds that we will converge toward the better approach. Bayes's theorem predicts that the Bayesians will win.

77. Fengyi Jin, et al., "Per-Contact Probability of HIV Transmission in Homosexual Men in Sydney in the Era of HAART," *AIDS*, 24, pp. 907–913, 2010. http://www.who.int/hiv/events/artprevention/jin_per.pdf.
78. Much of the research suggested that it was HIV-positive men who were driving the trend: most HIV-positive men would much prefer to have sex with other HIV-positive partners, particularly if they are not planning to use a condom. The advent of the Internet, as well as various types of support networks in the offline world, has made that much easier to do.
79. Larry Green, "Measles on Rise Nationwide; Chicago Worst Hit," *Los Angeles Times*, August 5, 1989. http://articles.latimes.com/1989-08-05/news/mn-469_1_chicago-health.
80. Justin Lessler et al., "Transmissibility of Swine Flu at Fort Dix, 1976," *Journal of the Royal Society Interface*, 4, no. 15, pp. 755–762, August 2007. http://rsif.royalsocietypublishing.org/content/4/15/755.full.
81. Ibid.
82. "Keep it sophisticatedly simple" was a phrase used by the late economist Arnold Zellner.
83. "Healthy Hand Washing Survey 2011," Bradley Corp. http://www.bradleycorp.com/handwashing/survey.jsp.
84. http://www.altpenis.com/penis_news/20060710032108data_trunc_sys.shtml.
85. "An Agent-Based Approach to HIV/AIDS Epidemic Modeling: A Case Study of Papua New Guinea, thesis, Massachusetts Institute of Technology, 2006. http://dspace.mit.edu/handle/1721.1/34528.
86. Shan Mei, et al., "Complex Agent Networks Explaining the HIV Epidemic Among Homosexual Men in Amsterdam," *Mathematics and Computers in Simulation*, 80, no. 5, January 2010. http://portal.acm.org/citation.cfm?id=1743988.
87. Donald G. McNeil Jr., "Predicting Flu with the Aid of (George) Washington," *New York Times*, May 3, 2009. http://www.nytimes.com/2009/05/04/health/04model.html?hp.
88. Michael A. Babyak, "What You See May Not Be What You Get: A Brief, Nontechnical Introduction to Overfitting in Regression-Type Models," *Statistical Corner, Psychosomatic Medicine*, 66 (2004), pp. 411–421.
89. Even if a prediction model is just a sort of thought experiment that is years away from producing useful results, it can still help us understand the scope of a problem. The Drake equation, a formula that provides a framework for predicting the number of intelligent extraterrestrial species in the galaxy, is not likely to yield highly useful and verifiable predictions in the span of our lifetimes—nor, probably, in the span of human civilization. The uncertainties are too great. Too many of its parameters are not known to within an order of magnitude; depending on which values you plug in, it can yield answers anywhere from that we are all alone in the universe to that there are billions and billions of extraterrestrial species. However, the Drake equation has nevertheless been a highly useful lens for astronomers to think about life, the universe, and everything.
90. George E. P. Box and Norman R. Draper, *Empirical Model-Building and Response Surfaces* (New York: Wiley, 1987), p. 424.
91. "Norbert Wiener," Wikiquote.org. http://en.wikiquote.org/wiki/Norbert_Wiener.

CHAPTER 8: LESS AND LESS AND LESS WRONG

1. Roland Lazenby, *The Show: The Inside Story of the Spectacular Los Angeles Lakers in the Words of Those Who Lived It* (New York: McGraw-Hill Professional, 2006).
2. Mark Heisler, "The Times' Rankings: Top to Bottom/NBA," *Los Angeles Times*, November 7, 1999.
3. Tom Spousta, "Pro Basketball: Trail Blazers Have Had Some Success Containing O'Neal," *New York Times*, May 20, 2000. http://www.nytimes.com/2000/05/20/sports/pro-basketball-trail-blazers-have-had-some-success-containing-o-neal.html?scp=2&sq=lakers+portland&st=nyt.
4. "Blazer Blowout Shows Need for 'Sheed," Associated Press; May 22, 2000. http://web.archive

.org/web/20041226093339/http://sportsmed.starwave.com/nba/2000/20000522/recap/porlal .html.

5. Tom Spousta, "Pro Basketball: Game 2 Was a Blur as Lakers Lost Focus," *New York Times*, May 24, 2000. http://www.nytimes.com/2000/05/24/sports/pro-basketball-game-2-was-a-blur -as-lakers-lost-focus.html?scp=3&sq=lakers+portland&st=nyt.
6. Tom Spousta, "Pro Basketball: Lakers Rally and Get Back on Track," *New York Times*, May 27, 2012. http://www.nytimes.com/2000/05/27/sports/pro-basketball-lakers-rally-and-get-back -on-track.html?scp=14&sq=lakers+portland&st=nyt.
7. Tom Spousta, "Pro Basketball: Everything Comes Up Roses for the Lakers," *New York Times*, May 29, 2000. http://www.nytimes.com/2000/05/29/sports/pro-basketball-everything-comes -up-roses-for-the-lakers.html?scp=16&sq=lakers+portland&st=nyt.
8. "Seventh Heaven: Blazers Send Series Back to L.A. for Game 7," Associated Press via *Sports Illustrated*, June 3, 2000. http://sportsillustrated.cnn.com/basketball/nba/2000/playoffs/ news/2000/06/02/lakers_blazers_gm6_ap/.
9. That is, $300,000 from winning his $200,000 bet on Portland at 3-to-2 odds, less the $80,000 that Voulgaris originally bet on the Lakers.
10. Tom Spousta, "Pro Basketball: Trail Blazers Follow Plan to the Bitter End," *New York Times*, June 7, 2000. http://www.nytimes.com/2000/06/05/sports/pro-basketball-trail-blazers-follow -plan-to-the-bitter-end.html?scp=28&sq=lakers+portland&st=nyt.
11. Per play-by-play data downloaded from Basketballvalue.com. http://basketballvalue.com/ downloads.php.
12. This is based on a logistic regression analysis I conducted of all games played in the 2009–2010 NBA regular season, where the independent variable is the score margin between the home team and the away team with fourteen minutes left to play in the game, and the dependent variable is whether or not the home team ultimately won. The regression model yields a value of .056 when the scoring margin is –16; that is, the home team has a 5.6 percent chance of victory when trailing by sixteen points, which translates into odds 17-to-1 against. I round down slightly to 15 to 1 because a team trailing by sixteen points at home will usually be inferior to its opponent, whereas the Lakers and Blazers were more evenly matched.
13. Voulgaris's odds of winning his bet at the start of the evening were about 50 percent: a 60 percent chance that the Lakers beat the Blazers in Game 7 multiplied by what I've estimated to be an 83 percent chance that the Lakers would beat the Pacers if advancing to the final. By that point in the game, the Lakers' odds of winning the Championship were down to about 5 percent: a 6 percent chance of coming back to beat the Blazers, multiplied by an 83 percent chance of beating the Pacers.
14. Miranda Hitti, "Testosterone Ups Home Field Advantage," *WebMD Health News*, June 21, 2006. http://www.webmd.com/fitness-exercise/news/20060621/testosterone-ups-home-field -advantage.
15. Most sports leagues hold their drafts in reverse order of finish: the team with the worst record is the first to pick. In the NBA, a sport where a single superstar talent can make an exceptional amount of difference, the league holds a draft lottery so as to discourage teams from tanking their games at the end of the season to improve their drafting position. Nevertheless, the worse a team does, the more Ping-Pong balls it gets in the lottery, and so teams will often play something other than their best basketball in these scenarios.
16. This asymmetry would not exist to the same extent if basketball teams were more focused on individual defensive statistics. But offense is relatively easy to measure, and defense is relatively hard; some teams don't even try to measure individual defensive performance at all. A player who scores a basket will therefore gain more market value than the man defending will lose by conceding one.
17. "2001–02 Cleveland Cavaliers Schedule and Results," Basketball-Reference.com. http://www .basketball-reference.com/teams/CLE/2002_games.html.
18. On average, a team will go either over or under the total five games in a row about five times

per season. That works out to 150 such streaks per season between the thirty NBA teams combined.

19. D. R. Bellhouse, "The Reverend Thomas Bayes FRS: A Biography to Celebrate the Tercentenary of His Birth," *Statistical Science*, 19, 1, pp. 3–43; 2004. http://www2.isye.gatech.edu/~brani/isyebayes/bank/bayesbiog.pdf.
20. Bayes may also have been an Arian, meaning someone who followed the teachings of the early Christian leader Arias and who regarded Jesus Christ as the divine son of God rather than (as most Christians then and now believe) a direct manifestation of God.
21. Thomas Bayes, "Divine Benevolence: Or an Attempt to Prove That the Principal End of the Divine Providence and Government Is the Happiness of His Creatures." http://archive.org/details/DivineBenevolenceOrAnAttemptToProveThatThe.
22. Ibid.
23. Ibid.
24. The Late Rev. Mr. Bayes, Communicated by Mr. Price, in a Letter to John Canton, M. A. and F. R. S., "An Essay Towards Solving a Problem in the Doctrine of Chances," *Philosophical Transactions of the Royal Society of London*, 53, pp. 370–418; 1763. http://www.stat.ucla.edu/history/essay.pdf.
25. Donald A. Gillies, "Was Bayes a Bayesian?," *Historia Mathematica*, 14, no. 4, pp. 325–346, November 1987. http://www.sciencedirect.com/science/article/pii/0315086087900656.
26. David Hume, "Cause and Effect" in *An Enquiry Concerning Human Understanding (1772)* (Hackett Publishing Company, 1993). http://www.marxists.org/reference/subject/philosophy/works/en/hume.htm.
27. Some Christians regard Bayesian probability as more compatible with their worldview. Under Bayes's theorem, if you assign a 100 percent prior probability to the hypothesis that a Christian God exists, then no amount of worldly evidence will shake you from that conviction. It is plausible that Bayes was aware of this property; in introducing Bayes's essay, Richard Price mentioned that he thought Bayes's theorem helped to confirm "the existence of the Diety."

    For further discussion, see Steve Bishop, "Christian Mathematicians—Bayes, God & Math: Thinking Christianly About Mathematics," *Education*, March 22, 2012. http://godandmath.com/2012/03/22/christian-mathematicians-bayes/.
28. "Fundamental Atheism," Free Atheist Church. https://sites.google.com/site/freeatheistchurch/fundamental-atheism.
29. Sharon Bertsch McGrayne, *The Theory That Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines, and Emerged Triumphant from Two Centuries of Controversy* (New Haven, CT: Yale University Press, Kindle edition), 427–436.
30. E. O. Lovett, "The Great Inequality of Jupiter and Saturn," *Astronomical Journal*, 15, 351 (1895), pp. 113–127.
31. McGrayne, *The Theory That Would Not Die*, Kindle location 19.
32. Pierre-Simon Laplace, "A Philosophical Essay on Probabilities" (1902), pp. 6–8.
33. Bret Schulte, "How Common Are Cheating Spouses?" *U.S. News & World Report*, March 27, 2008. http://www.usnews.com/news/national/articles/2008/03/27/how-common-are-cheating-spouses.
34. "Breast Cancer Risk by Age," Breast Cancer Centers for Disease Control and Prevention, last updated August 13, 2010. http://www.cdc.gov/cancer/breast/statistics/age.htm.
35. "Understanding Breast Exam Results—False Negative–False Positive Results," RealAge.com. http://www.realage.com/womens-health/breast-exam-results.
36. S. Eva Singletary, Geoffrey L. Robb, and Gabriel N. Hortobagyi, "Advanced Therapy of Breast Disease," B. C. Decker, May 30, 2004.
37. Gina Kolata, "Panel Urges Mammograms at 50, Not 40," *New York Times*, November 16, 2009. http://www.nytimes.com/2009/11/17/health/17cancer.html.
38. Dan M. Kahan, et al., "The Polarizing Impact of Science Literacy and Numeracy on Per-

ceived Climate Change Risks," *Nature Climate Change*, May 27, 2012. See Supplementary Information: http://www.nature.com/nclimate/journal/vaop/ncurrent/extref/nclimate1547-s1.pdf.

39. Twenty-five thousand days prior to September 11, 2001, would take us back to 1942.
40. John P. A. Ioannidis, "Why Most Published Research Findings Are False," *PLOS Medicine*, 2, e124, August 2005. http://www.plosmedicine.org/article/info:doi/10.1371/journal.pmed.0020124.
41. Brian Owens, "Reliability of 'New Drug Target' Claims Called into Question," *NewsBlog, Nature*, September 5, 2011. http://blogs.nature.com/news/2011/09/reliability_of_new_drug_target.html.
42. McGrayne, *The Theory That Would Not Die*, Kindle location 46.
43. Paul D. Stolley, "When Genius Errs: R. A. Fisher and the Lung Cancer Controversy," *American Journal of Epidemiology*, 133, 5, 1991. http://www.epidemiology.ch/history/PDF%20bg/Stolley%20PD%201991%20when%20genius%20errs%20-%20RA%20fisher%20and%20the%20lung%20cancer.pdf.
44. Alan Agresti and David B. Hitchcock, "Bayesian Inference for Categorical Data Analysis," *Statistical Methods & Applications*, 14 (2005), pp. 297–330. http://www.stat.ufl.edu/~aa/articles/agresti_hitchcock_2005.pdf.
45. John Aldrich, "R. A. Fisher on Bayes and Bayes' Theorem," *Bayesian Analysis*, 3, no. 1 (2008), pp. 161–170. http://ba.stat.cmu.edu/journal/2008/vol03/issue01/aldrich.pdf.
46. McGrayne, *The Theory That Would Not Die*, Kindle location 48.
47. Tore Schweder, "Fisherian or Bayesian Methods of Integrating Diverse Statistical Information?" *Fisheries Research*, 37, 1–3 (August 1998), pp. 61–75. http://www.sciencedirect.com/science/article/pii/S0165783698001271.
48. 2008 New Hampshire Democratic Primary polls via RealClearPolitics.com. http://www.realclearpolitics.com/epolls/2008/president/nh/new_hampshire_democratic_primary-194.html.
49. Nate Silver, "Rasmussen Polls Were Biased and Inaccurate; Quinnipiac, SurveyUSA Performed Strongly," *FiveThirtyEight, New York Times*, November 4, 2010. http://fivethirtyeight.blogs.nytimes.com/2010/11/04/rasmussen-polls-were-biased-and-inaccurate-quinnipiac-surveyusa-performed-strongly/.
50. R. A. Grant and T. Halliday, "Predicting the Unpredictable: Evidence of Pre-Seismic Anticipatory Behaviour in the Common Toad," *Journal of Zoology*, 700, January 25, 2010. http://image.guardian.co.uk/sys-files/Environment/documents/2010/03/30/toads.pdf.
51. "Hate Group Formation Associated with Big-Box Stores," *ScienceNewsline.com*, April 11, 2012. http://www.sciencenewsline.com/psychology/2012041121000031.html.
52. Aldrich, "R. A. Fisher on Bayes and Bayes' Theorem."
53. McGrayne, *The Theory That Would Not Die*, Kindle location 111.
54. Sir Ronald A. Fisher, "Smoking: The Cancer Controversy," Oliver and Boyd. http://www.york.ac.uk/depts/maths/histstat/smoking.htm.
55. Jean Marston, "Smoking Gun," *NewScientist*, no. 2646, March 8, 2008. http://www.newscientist.com/article/mg19726460.900-smoking-gun.html.
56. McGrayne, *The Theory That Would Not Die*, Kindle location 113.
57. Stolley, "When Genius Errs."
58. Ibid.
59. Jo Tuckman and Robert Booth, "Four-Year-Old Could Hold Key in Search for Source of Swine Flu Outbreak," *The Guardian*; April 27, 2009. http://www.guardian.co.uk/world/2009/apr/27/swine-flu-search-outbreak-source
60. McGrayne, *The Theory That Would Not Die*, Kindle location 7.
61. Raymond S. Nickerson, "Null Hypothesis Significance Testing: A Review of an Old and Continuing Controversy," *Psychological Methods*, 5, 2 (2000), pp. 241–301. http://203.64.159.11/richman/plogxx/gallery/17/%E9%AB%98%E7%B5%B1%E5%A0%B1%E5%91%8A.pdf.
62. Andrew Gelman and Cosma Tohilla Shalizi, "Philosophy and the Practice of Bayesian Sta-

tistics," *British Journal of Mathematical and Statistical Psychology*, pp. 1–31, January 11, 2012. http://www.stat.columbia.edu/~gelman/research/published/philosophy.pdf.

63. Although there are several different formulations of the steps in the scientific method, this version is mostly drawn from "APPENDIX E: Introduction to the Scientific Method," University of Rochester. http://teacher.pas.rochester.edu/phy_labs/appendixe/appendixe.html.
64. Thomas S. Kuhn, *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press, Kindle edition).
65. Jacob Cohen, "The Earth Is Round (p<.05)," *American Psychologist*, 49, 12 (December 1994), pp. 997–1003. http://ist-socrates.berkeley.edu/~maccoun/PP279_Cohen1.pdf.
66. Jeff Gill, "The Insignificance of Null Hypothesis Significance Testing," *Political Research Quarterly*, 52, 3 (September 1999), pp. 647–674. http://www.artsci.wustl.edu/~jgill/papers/hypo.pdf.
67. David R. Anderson, Kenneth P. Burnham, and William L. Thompson, "Null Hypothesis Testing: Problems, Prevalence, and an Alternative," *Journal of Wildlife Management*, 64, 4 (2000), pp. 912–923. http://cat.inist.fr/%3FaModele%3DafficheN%26cpsidt%3D792848.
68. William M. Briggs, "It Is Time to Stop Teaching Frequentism to Non-Statisticians," arXiv.org, January 13, 2012. http://arxiv.org/pdf/1201.2590.pdf.
69. David H. Krantz, "The Null Hypothesis Testing Controversy in Psychology," *Journal of the American Statistical Association*, 44, no. 448 (December 1999). http://www.jstor.org/discover/10.2307/2669949?uid=3739832&uid=2&uid=4&uid=3739256&sid=47698905120317.

CHAPTER 9. RAGE AGAINST THE MACHINES

1. "Poe Invents the Modern Detective Story," National Historic Site Philadelphia, National Park Service, U.S. Department of the Interior. http://www.nps.gov/edal/forteachers/upload/detective.pdf.
2. Nick Eaton, "Gallup: Bill Gates Is America's Fifth-Most Admired Man," *Seattle Post-Intelligencer*, December 27, 2010. http://blog.seattlepi.com/microsoft/2010/12/27/gallup-bill-gates-is-americas-fifth-most-admired-man/.
3. Joann Pan, "Apple Tops Fortune's 'Most Admired' List for Fifth Year in a Row," *Mashable*, March 2, 2012. http://mashable.com/2012/03/02/apple-tops-fortunes-most-admired-list-five-years-straight-video/.
4. David Kravets, "Stock-Picking Robot 'Marl' Is a Fraud, SEC Says," Threat Level, *Wired*, April 23, 2012. http://www.wired.com/threatlevel/2012/04/stock-picking-robot/.
5. "What Is the Stock Trading Robot 'MARL'?," Squidoo.com. http://www.squidoo.com/StockTradingRobotMARL.
6. *Philadelphia Inquirer*, "Computer Predicts Odds of Life, Death," *Orlando Sentinel*, July 9, 1992. http://articles.orlandosentinel.com/1992-07-09/news/9207090066_1_apache-system-critical-care-critical-care.
7. Nick Montfort, *Twisty Little Passages: An Approach to Interactive Fiction* (Boston: MIT Press, 2005), p. 76.
8. Claude E. Shannon, "Programming a Computer for Playing Chess," *Philosophical Magazine*, Series 7, 41, 314, March 1950. http://archive.computerhistory.org/projects/ chess/related_materials/software/2-0%20and%202-1.Programming_a_computer_for_playing_chess.shannon/2-0%20and%202-1.Programming_a_computer_for_playing_chess.shannon.062303002.pdf.
9. William G. Chase and Herbert A. Simon, "The Mind's Eye in Chess" in *Visual Information Processing* (New York: Academic Press, 1973).
10. Douglas Harper, *Online Etymology Dictionary*. http://www.etymonline.com/index.php?term=eureka.
11. Amos Tversky and Daniel Kahneman, "Judgement Under Uncertainty: Heuristics and Biases," *Science*, 185 (September 27, 1974), pp. 1124–1131. http://www.econ.yale.edu/~nordhaus/homepage/documents/tversky_kahn_science.pdf.