

# THE BLACK SWAN

The Impact of the Highly Improbable

Nassim Nicholas Taleb



RANDOM HOUSE  
TRADE PAPERBACKS







# THE BLACK SWAN

*The Impact of the Highly Improbable*

NASSIM NICHOLAS TALEB



RANDOM HOUSE TRADE PAPERBACKS  
NEW YORK



























# UMBERTO ECO'S ANTILIBRARY, OR HOW WE SEEK VALIDATION

The author Umberto Eco believes that there are two types of people: those who are insightful and considerate, and those who are not. These two categories correspond to the two main types of visitors into bookstores: those who buy with their Sophsophore professors' dollars, Eco, what library you have! How many of these books have you read? and the others—a very small minority—who go the path that a private library is not an ego-booster, dependence, but a research tool. Read books are far less valuable than unread ones. The library should contain as much of what you do not know as your financial means, mortgage rates, and the currently tight real estate market allow you to put there. You will accumulate more knowledge and more books as you grow older, and the growing number of unread books on the shelves will look at you menacingly. Indeed, the more you know, the larger the mass of unread books. Let us call this collection of unread books an antilibrary.

We tend to treat our knowledge as personal property to be protected and defended. It is an ingrained trait that allows us to rise in the pecking order. So this tendency to offend Eco's library sensibly by focusing on the unknown is a human bias that extends to our mental operations. People don't walk around with anti-resumes telling you what they have not studied or experienced (it's the job of their competitors to do many, but it would be nice if they did). Just as we tend to stand library books on its head, we will work on standing knowledge itself on its head. Note that the Black Swan comes from our misunderstanding of the likelihood of surprises, those unread books, because we take what we know a little too seriously.

Let us call an antischolar—someone who focuses on the unread books, and makes an attempt not to treat his knowledge as a treasure, or even a possession, or even a self-esteem enhancement device—a skeptical empiricist.

The chapters in this section address the question of how we humans deal with knowledge. In our reference to the anecdotal over the empirical, [Chapter 1](#) presents the Black Swan as ground truth, and in my own discussion, I will make a central distinction between the two varieties of randomness. [Chapter 3](#)

[Chapter 4](#) finally returns to the Black Swan problem in its original form: how we tend to generalize from what we see, then present the three facets of the same Black Swan problem: a) The error of confirmation, or how we are likely to underestimate some the other part of the library, the tendency to look at what confirms our knowledge; b) the narrative fallacy, or how we are likely to overestimate the importance of our own personal history; c) the heretic fallacy, or how we are likely to underestimate the importance of the other parts of the library.

[Chapter 5](#), the narrative fallacy, [Chapter 6](#), the heretic fallacy, [Chapter 7](#), the error of confirmation, [Chapter 8](#), the narrative fallacy, [Chapter 9](#), the heretic fallacy.



























## Extremistan

## Scalable

With zero dependency or type I randomness.

The most “typical” is either giant or dwarf, i.e., there is no typical number.

## Whiner-like moment effects

Takes a long time to get off.

More likely to be found in our modern environment.

## Unscalable in the Black Swan

There are no physical constraints on what a number can be.

Corresponds to numbers very small.

Dominated by extreme winner take-all inequality.

Value will be determined by a small number of extreme events.

It takes a long time to know what's going on.

## Typology of the accidental

Hard to predict from past information.

History makes jumps.

The distribution is either Mandelbrotian (gray Swans) tractable scientifically, or totally unpredictable Black Swans.

\* What I call “probability distribution” here is the model used to calculate the odds of different events, how they are distributed. When I say that an event is distributed according to the “bell curve,” I mean that the Gaussian bell curve (after C. F. Gauss; more on him later) can help provide probabilities of various occurrences.

Mandelbrot does not always imply Black Swans. Some events can be rare and consequential, but somewhat predictable, according to those who are prepared for them and have the tools to understand them (such as listening to meteorologists, economists, and historians of the bell curve variety). They are near-Black Swans—they are somewhat predictable—knowing about their existence should lower your surprise that events are rare but expected. I call this special case of “gray” swans Mandelbrotian randomness. This category encompasses the randomness that produces phenomena commonly known by terms such as: seismic, accident, power laws, Pareto-Zipf laws, Valley law, Power-law processes, Zipfian law, and fractal law, and we will leave them aside for now since they will be covered in some depth in Part Three. They are scalable according to the logic of this chapter, but you can know a little more about how they scale since they share much with the laws of nature.

You can still experience severe Black Swans in Mandelbrotian, though not easily. How? You may forget that something is random, think that it is deterministic, then have a surprise. Or you can ignore and miss on a source of uncertainty, whether due to a lack of imagination—most Black Swans result from this “lumping” disease, which will discuss in [Chapter 9](#).\*

This has been a “theory” overview of the central distinction of this book, offering a link to distinguish between what can belong in Mandelbrotian and what belongs in Extremistan. I said that I will get into a more thorough examination in Part Three, so let us focus on epistemology for now and see how the distinction affects our knowledge.

\* To those readers who Googled Yevgenia Katsenelenbaum, I am sorry to say that she is (officially) fictional character.

\* I emphasize possible because the chance of these occurrences is typically in the order of one in several

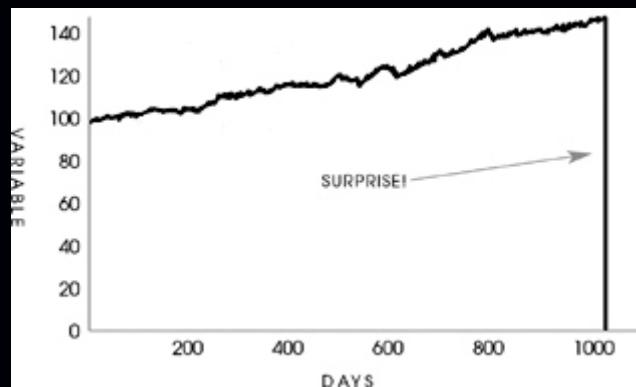




not wings you can fly. Consider the case of the increasingly integrated German Jews in the 1930s—our description in [Chapter 1](#) of how the population of Lebanon got lulled into a false sense of security by the appearance of greater kindness and tolerance.

Let us go one step further and consider induction's most worrisome aspect, learning backward. Consider the history's experience may have, rather than no value, a negative value. It learned from observation, as we all wished to do they after all, this is what is believed to be the scientific method. Its confidence increased as the number of friendly feelings grew, and it felt increasingly safe even though the slaughter was more and more imminent. Consider that the feeling of safety reached its maximum when the risk was at the highest but the problem is even more general than that, it strikes at the nature of empirical knowledge itself. Something has worked in the past, until—well, it unexpectedly no longer does, and what we have learned from the past turns out to be at best irrelevant or worse, it was actively misleading.

**FIGURE 1. ONE THOUSAND AND ONE DAYS OF HISTORY**



A turkey before and after Thanksgiving: the history of a process over 1,000 days tells you nothing about what is to happen next. The naive projection of the future from the past can be applied to anything.

**Figure 1** provides the mathematical case of the problem of induction as articulated in [Table 1](#). You observe a certain variable in one thousand days. It could be anything (births, homicides, transformations, stock price, pressure gauge, your personal income, a given stock, the interest rate, etc.). In Sunday Abundance of a specific Greek Orthodox church. You subsequently derive, solely from past data, a few conclusions concerning the properties of the pattern with projections for the next thousand, even five thousand, days. On the one thousand and first day—boom! A big change takes place that is completely unprepared for by the past. Consider the surprise of the Great War. After the Napoleonic conflicts, the world had experienced a period of peace that would lead any observer to believe in the disappearance of severely destructive conflicts. Yet surprise! It turned out to be the deadliest conflict, up until then, in the history of mankind.

Note that after the event you start predicting the possibility of other outliers happening locally, that is, in the process you were just surprised by—but not elsewhere. After the stock market crash of 1987, well, American traders braced for another one every October—not taking into account that there was no precedent for the first one. We worry too late—ex post. Making a naive observation of the past as something definitive representative of the future is the one and only cause of our inability to understand the Black Swan.

It would appear to a sober statistician—i.e., one of those writers and scholars who fill up their books with citations from some dead authority—that so phrased by Hobbes, “From like antecedents, flow like consequents.” Those who believe in the unconditional benefits of past experience should consider this pearl of wisdom allegedly voiced by a famous ship’s captain:

But in all my experience I have never been in any accident, or any misfortune speaking about I have been on one vessel in distress in all my years at sea. I never saw a wreck and never have been wrecked nor was I ever in any predicament that threatened to end in disaster of any sort.

—E. J. Smith, 1907, Captain, RMS *Titanic*

Captain Smith’s ship sank in 1912 in what became the most talked-about shipwreck in history. \*

#### Trained to Be Dull

Similarly, think of a bank chairman whose institution makes steady profits over a long time, only to lose everything in a single reversal of fortune. Traditionally, bankers of the lending variety have been past-shapers, clean-shavers, and dress in possibly the most comforting and boring manner in dark suits, white shirts, and red ties. Indeed, for their lending business, banks hire dull people and train them to be even more dull. But this is to show, if they look conservative, it is because their loans only go bust on rare, very rare, occasions. There is no way to gauge the effectiveness of their lending activity by observing it over a day, a week, a month, or... even a century. In the summer of 1902, large American banks lost close to all their past earnings cumulatively, about everything they ever made in the history of American banking—everything. They had been lending to South and Central American countries that all defaulted at the same time—an event of an exceptional nature. So it took just one summer to figure out that this was a sucker’s business and that all their earnings came from a single risky game. All that while the bankers, not everyone, especially themselves, into believing that they were “conservative.” They are not conservative, just phenomenally skilled at self-deception by burying the possibility





























































He has no incentive to do so. The losses of the person hurt by the side effects will go after the doctor. His attack deep, while the lives saved by the drug might not be accounted for anywhere.

A life saved is a statistic, a person hurt is an anecdote. Statistics are invisible; anecdotes are salient. Likewise, the risk of a Black Swan is invisible.\*



Giacomo Casanova, aka. Jacques Casanova de Seingalt, the famous seducer and legendary seducer did not look quite like James Bond.

#### THE TEFLON-STYLE PROTECTION OF GIACOMO CASANOVA

This brings us to gravest of all manifestations of silent evidence, the illusion of stability. The ones known or perception of the risks we incurred in the past, particularly for those of us who were lucky to have survived them. Your life came under a serious threat but, having survived it, you retrospectively underestimate how risky the situation actually was.

The adventurer Giacomo Casanova, later self-styled Jacques Casanova de Seingalt, the famous seducer and legendary seducer of women, seems to have had a Teflon-style trait that would cause envy on the part of the most resolute of Metal fans: misfortune did not stick to him. Casanova, while known for his seductions, viewed himself as some sort of a scholar. He signed at literary fame with his twelve-volume *History of My Life*, written in bad (charmingly bad) French, in addition to the extremely useful lessons on how to become a seducer. The *History* provides an engrossing account of a succession of reversals of fortune. Casanova felt that every time he got into difficulties, his lucky star, his *étoile*, would pull him out of trouble. After things got bad for him, they somehow recovered by some invisible hand, and he was led to believe that it was his intense propensity to recover from hardships by turning every time into a new opportunity. He would somehow meet someone in extreme who offered him a financial transaction, a new patron that he had not betrayed in the past, or someone generous enough who with a weak enough memory to forget past betrayals. Could Casanova have been selected by destiny to bounce back from all hardships?

Not necessarily. Consider the following: of all the colonial adventurers who have lived on our planet, many were occasionally crushed, and a few did bounce back repeatedly. It is those who survive who will tend to believe that they are indestructible; they will have a long and interesting enough experience to write books about it. Until, of course,

Actually, adventurers who feel singled out by destiny abound, simply because there are plenty of adventurers, and we do not hear the stories of those down on their luck. As I started writing this chapter, I recalled a conversation with a woman about her Tambourin René, the son of a civil servant, who managed through a few financial transactions to catapult himself into the life of a character in a novel, with handmade shoes, Cuban cigars, collector cars, and so on. The French have a word for this: *tambourin*, which means a mixture of extravagantly boastful with speculator, who risk takes all the while keeping considerable personal charm, a word that does not seem to be available in Anglo-Saxon cultures. The *tambourin* was spending his money very quickly, and as we were having the conversation about his fate (he was going to marry him, after all), she explained to me that he was undergoing slightly difficult times but that there was no need to worry since he always come back with a vengeance. That was a few years ago. Out of curiosity, I have just tracked him down trying to do so tactfully; he has not recovered fully from his latest bout of failure. He also dropped out of the scene and is no longer to be found among other *tambourins*.

How does this relate to the dynamics of history? Consider what is generally called the resilience of New York City. For seemingly inscrutable reasons, every time it gets close to the brink of disaster, the city manages to pull back and recover. Some people truly believe that this is an internal property of New York City. The following quote is from a *New York Times* article:

Which is why New York still needs Samuel H. C., an economist who turns 77 today. Mr. C. studied New York City through half a century of booms and busts. ... "We have a record of going through tough times and coming back stronger than ever," he said.

Now run the idea in reverse: think of cities as little Giacomo Casanovas, or as rats in my laboratory. As we put

















# WE JUST CAN'T PREDICT

When I ask people to name three major unanticipated technologies that have impacted our world today, they usually propose the computer, the Internet, and the laser. All three were uncharted, unpredicted, and unanticipated upon their discovery, and remained unanticipated well after their initial use. They were *Black Swans*. Of course, we have the retrospective illusion of their lurking in some master plan. You can create your own lists with similar results, whether you use political events, wars, or medical discoveries.

You would expect our record of prediction to be horrible; the world is far, far more complicated than we think, which is not a problem, except when most of us don't know it. We tend to tunnel while looking into the future, making it business as usual. Black Swans-free, when in fact there is nothing *new* about the future, it is not a *Historic category*.

We have seen how good we are at narrating backward, at inventing stories that convince us that we understand the past. For many, people, knowledge has the remarkable power of producing confidence instead of measurable aptitude. Another problem: the focus on the (inconsequently) regular, the *Histotification* that makes the forecasting "inside the box."

I find it scandalous that in spite of the empirical record we continue to project into the future as if we were good at it using tools and methods that exclude rare events. Prediction is firmly institutionalized in our world. We are suckers for those who help us navigate uncertainty, whether they fortunetell in the "well-published" (that is, academics or civil servants) using phony mathematics,

## From Yogi Berra to Henri Poincaré

The great baseball coach Yogi Berra has a saying, "It is tough to make predictions, especially about the future," while he did not produce the writings that would allow him to be considered a philosopher in spite of his wisdom and intellectual abilities. Berra can claim to know something about randomness. He was a practitioner of uncertainty, and, as a baseball player and coach, regularly faced random outcomes, and had to face them results deep into his bones.

In fact, Yogi Berra is not the only thinker who thought about how much of the future lies beyond our abilities. Many less popular, less witty, but no less competent thinkers than he have examined our inherent limitations in this regard. From the philosopher Jacques Hadamard and Henri Poincaré (commonly described as mathematicians) to the philosopher Rudolf von Blythe (commonly described, alas, as an economist), to the philosopher Karl Popper (commonly known as a philosopher), we can safely call this the *Berra-Hadamard-Poincaré-Blythe-Popper* consensus, when puts structural, built-in limits to the enterprise of predicting. "The future ain't what it used to be," Berra later said.\* He seems to have been right; the gains in our ability to model (and predict) the world may be dwarfed by the increases in its complexity—imposing a greater and greater role for the unpredictable. The larger the role of the *Black Swan*, the harder it will be for us to predict. Sorry.

Before going into the limits of prediction, we will discuss our track record in forecasting and the relation between gains in knowledge and the offsetting gains in confidence.

\* Note that these sayings attributed to Yogi Berra might be apocryphal—it was the physicist Niels Bohr who came up with the first one, and plenty of others came up with the second. These sayings remain, however, quintessential *Black Swans*.























happened in their jobs, who forecast simply because "that's my job," knowing very well that their forecasts are not what I would call ethical. What they do is no different from robbing, but less simply because "it's my job."

Anyone who causes harm by forecasting should be treated as either a fool or a bad one. Some forecasts do more damage to society than criminals. Please, don't drive a school bus blindfolded.

As JPK

At New York's JFK airport you can find gigantic newsstands with walls full of magazines. They are usually manned by a very polite family from the Indian subcontinent just the parents, the children are in medical school. These walls present you with the entire range of what an "informed" person needs in order to know what's going on. I wonder how long it would take to read every single one of these magazines, excluding fishing and motorcycle periodicals (but including the gossip magazines—you might as well have some fun). How a lifetime? An entire lifetime?



Consequently, The Fortune-Teller. We have always been suckers for those who tell us about the future. In this picture the fortune-teller is stealing the woman's ring.

Sadly, all this knowledge would not help the reader to forecast what is to happen tomorrow. Actually, it might decrease his ability to forecast.

There is another aspect to the problem of prediction: its inherent limitations, those that have little to do with human nature, but instead arise from the very nature of information itself. I have said that the Black Swan has three attributes: unpredictability, consequences, and retrospective explainability. Let us examine the third one briefly.

\* The book you have in your hands is approximately and "unexpectedly" fifteen months late.

\* While forecast errors have always been令人惊讶的, commodity prices have been a growing concern. Consider this: 1973 forecast by U.S. oiliers (signed by the U.S. Secretaries of the Treasury, State, Interior, and Defense). The standard price of foreign crude oil by 1990 may well decline and will in any event not experience a sustained increase. Oil prices were up until by 1990. I just wonder if current forecasters lack imagination or if they are intentionally ignoring forecast errors.

Please note this additional observation: since high oil prices are marking up their inventories, oil companies are making record bucks and oil executives are getting huge bonuses because they did a good job—so they brought profits by causing the rise of oil prices.

\* I leave the reader an answer concerning Catherine's river count. She had only twelve.







we frequently witness the proliferation of similar proofs coming out of nowhere with accusations of leakage and plagiarism. There may be no plagiarism, the information that the solution exists is itself a big piece of the solution.

By the same logic, we are not easily able to conceive of future inventions. If we were, they would have already been invented. On the day when we are able to foresee inventions we will be living in a state where everything conceivable has been invented. Our own condition brings to mind the apocryphal story from 1899 when the head of the U.S. patent office resigned because he deemed that there was nothing left to discover—except that on that day the resignation would be justified.\*

François was not the first to go after the limit to our knowledge. In Germany, in the late nineteenth century, Eugène Charles Raymond claimed that ignoramus et ignorabimus—we are ignorant and will remain so. Somehow his ideas were not popular. But not before causing a reaction, the mathematician David Hilbert set to deny him by drawing a list of problems that mathematicians would need to solve over the next century.

Even if Raymond was wrong, we are not even good at understanding the unknowable. Consider the statements we make about things that we will never come to know—we confidently underestimate what knowledge we may acquire in the future. Auguste Comte, the founder of the school of positivism, which is usually defined as aiming at the scientization of everything in sight, declared that mankind would never know the secret of the chemical composition of the two stars. But as Charles Sanders Peirce reported, "This was scarcely dry upon the printed page before the spectroscope was discovered and that which he had deemed absolutely unknowable was well on the way of being ascertained." Comte's own predictions, concerning what we would come to learn about the workings of society, were grossly—and dangerously—overstated. He assumed that society was like a clock that would yield its secrets to us. It suffices to remember here Peirce's requires knowing about technologies that will be discovered in the future, but that very knowledge would almost automatically allow us to start developing those technologies right away. Ergo, we do not know what we will know.

Some might say that the argument is reduced since obvious that we always think that we have reached definitive knowledge but don't notice that those past societies we laugh at also thought the same way. The argument is true, so why don't we take it into account? The answer lies in a pathology of human nature. Remember the pathological discussions of economists in the formation of walls in the prison that we see from in others and from ourselves. That is what happened to the mathematical economist machine.



Mathematician Henri Poincaré, shown here seated in his study in the town of Gouville. Courtesy of Université Clermont Auvergne.

### THE NTH BILLIARD BALL

Henri Poincaré, in spite of his fame, is regularly considered to be an undervalued scientific thinker, given that he died in a century for some of his ideas to be appreciated. He was perhaps the one great pure mathematician of possibly the second, a mathematical thinker. Every time I see a T-shirt bearing the picture of the modern icon Albert Einstein, I cannot help thinking of Poincaré, Einstein is worthy of our reverence, but he has deserved many others. That is why the name of our conference is appropriate to him.

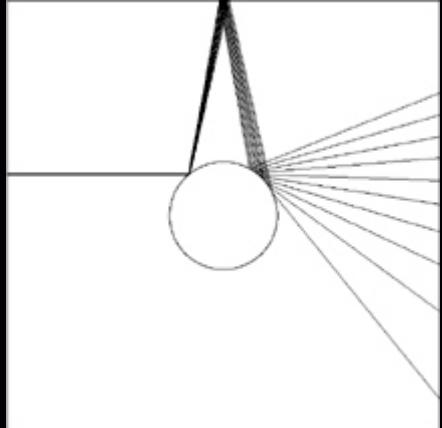
### Third Republic-Style Decorum

Again, Poincaré is in a class by himself. Recall my father recommending Poincaré's essays, not just for their scientific content, but for the quality of his French prose. The grand master wrote these articles as polished articles and composed them like extemporaneous speeches. As in every masterpiece, you see a mixture of repetitions, digressions, everything is the too strict with a prepackaged mind would condemn—but these make his text even more readable owing to an intrinsic consistency of thought.

Poincaré became a prolific essayist in his thirties. He seemed in a hurry and died prematurely at fifty-eight. He was in such a rush that he did not bother correcting typos and grammatical errors in his text, even after spotting them, since he found doing so a gross misuse of his time. They no longer make penises like that—or they no longer let them write in their own way.

Poincaré's reputation as a thinker ended rapidly after his death. His idea that contains us back almost a century





One of the readers of a draft of this book, Lewis Carroll, was very close to the point of becoming a mathematician, but he, like the second source quoted in the first section, never made it. In the mathematical version of the analysis developed, every additional source will be fully integrated. This causes a source-multiplicative effect where the error grows out of proportionality.

Excessive forecasting difficulty comes from complicating the mechanics, even so slightly. Our world, unfortunately, is far more complicated than the three-body problem. It contains far more than three objects. We are dealing with what is now called a dynamical system—and the world, we will see, is nothing but such a dynamical system.

Think of the difficulty in forecasting in terms of branches growing out of a tree: at every fork we have a multiplication of new branches. To get some intuition about these nonlinear, multiplicative effects, let's start with a well-known story about the chessboard. The inventor of the chessboard requested the following compensation: one grain of rice for the first square, two for the second, four for the third, eight, then sixteen, and so on, covering every one, square after square. The king granted the request, thinking that the inventor was asking for a paltry—but he soon realized that he was outsmarted. The amount of rice exceeded all possible grain reserves!

The mathematical difficulty leading to the need for greater and greater precision in assumptions can be illustrated with the following simple exercise concerning the prediction of the movements of billiard balls on a table. I use the example as computed by the mathematician Michael Berry. If you know a set of basic parameters concerning the ball at rest, compute the resilience of the table (this elementary), and compute the strength of the impact, then it is rather easy to predict what would happen at the first hit. The second impact becomes more complicated, but possible; you need to be more careful about your knowledge of the initial states, and more precision is called for. The problem is that to correctly compute the sixth impact, you need to take into account the gravitational pull of someone standing next to the table (modestly, her computations use a weight of less than 150 pounds). And to compute the fifteenth impact, every single elementary particle of the universe needs to be present in your assumption! An electron at the edge of the universe, separated from us by 10 billion light-years, must figure in the calculations, since it exerts a meaningful effect on the outcome. Now consider the additional burden of having to incorporate predictions about where these variables will be in the future. Forecasting the motion of a billiard ball on a pool table requires knowledge of the dynamics of the entire universe, down to every single atom. We can easily predict the movements of large objects like planets (though not too far into the future), but the smaller entities can be difficult to figure out—and there are so many more of them.

Note that this billiard-ball story assumes a plain and simple world; it does not even take into account these more social matters, possibly endowed with free will. Billiard balls do not have a mind of their own. Nor does our example take into account relativity and quantum effects. Nor did we use the notion often invoked by physicists, called the “uncertainty principle.” We are not concerned with the limitations of the precision in measurements done at the subatomic level. We are just dealing with billiard balls!

In a dynamical system, where you are considering more than a ball on its own, where trajectories may depend on one another, the ability to project into the future is not just reduced, but is subjected to a fundamental limitation. Poincaré proposed that we can only work with qualitative matters—some property of systems can be discussed, but not computed. You can think rigorously, but you cannot use numbers. Poincaré even invented a field for this analysis in situ, now part of topology. Prediction and forecasting are a more complicated business than is commonly accepted, but it takes someone who knows mathematics to understand that. To accept it takes both understanding and courage.

In the 1960s the MIT meteorologist Edward Lorenz rediscovered Poincaré's results on his own—once again, by accident. He was producing a computer model of weather dynamics and, before a simulation that predicted a weather system a few days ahead. Later he tried to repeat the same simulation with the exact same model and what he thought were the same input parameters, but he got wildly different results. He initially attributed these differences to a computer bug or a calculation error. Computers then were heavier and slower machines back then, no resemblance to what we have today, so users were severely constrained by time. Lorenz subsequently





individual will perform a unique set of actions in specified circumstances. There is one and only one answer to the question of how "rational" people satisfying their best interests would act. Rational actors must be coherent; they cannot prefer apples to oranges, oranges to pears, then pears to apples. If they did, then it would be difficult to generalize their behavior. It would also be difficult to predict their behavior in time.

In orthodox economics, rationality became a straitjacket. Platonized economists ignored the fact that people might prefer to do something other than maximize their economic interests. This led to mathematical techniques such as "maximization," or "optimization," on which Paul Samuelson built much of his work. Optimization consists in finding the mathematically optimal policy that an economic agent could pursue. For instance, what is the "optimal" quantity you should allocate to work? It involves complicated mathematics and this raises a barrier to entry by non-mathematically trained scholars. I would not be the first to say that this optimization set back social science by reducing it from the intellectual and reflective discipline that it was becoming to an attempt at an "exact science." By "exact science," I mean a secondary engineering problem for those who want to pretend that they are in the physics department—so-called physisics—in other words, an intellectual fraud.

Optimization is a case of sterile modeling that we will discuss further in [Chapter 17](#). It had no practical nor even theoretical use, and so it became principally a competition for academic positions, a way to make people compete with mathematical models. It kept Platonized economists out of the real world questions at night. The tragedy is that Paul Samuelson, a quick mind, is said to be one of the most intelligent scholars of his generation. This was clearly a case of very badly invested intelligence. Characteristically, Samuelson intimidated those who questioned his techniques with the statement, "Those who can, do science; others do methodology." If you know math, you could do science. This is reminiscent of psychoanalysts who silence their critics by accusing them of having conflicts with their fathers. And it bears out that it was Samuelson and most of his followers who did not know much math, or did not know how to use what math they knew, how to apply it to reality. They only knew enough math to be blinded by it.

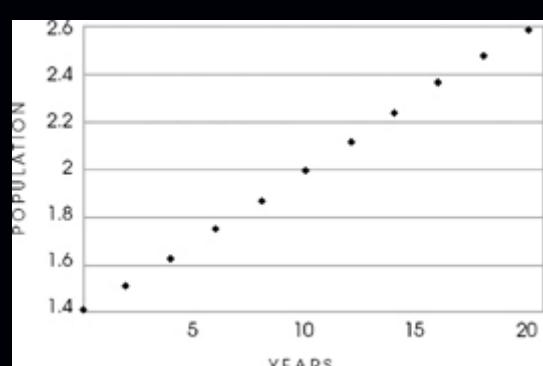
Frequently, before the proliferation of empirically blind (but severe) interesting work has been begun by free thinkers, the likes of J. M. Keynes, Friedrich Hayek, and the great Benoit Mandelbrot, all of whom were displaced because they moved economics away from the precision of second-rate physics. Very well. One great undiscovered thinker is G. C. Shackle, now almost completely obscure, who introduced the notion of "unknowledge," that is, the unred books in Umberto Eco's library. It is unusual to see Shackle's work mentioned at all, and I had to buy his books from secondhand dealers in London.

Lessons of empirical psychologists of the heuristics and biases school have shown that the model of rational behavior under uncertainty is not just grossly inaccurate but plain wrong as a description of reality. Their results also baffle Platonized economists because they reveal that there are several ways to be irrational. Tolstoy said that happy families are all alike, while each unhappy one is unhappy in its own way. People have been shown to make choices equivalent to preferring apples to oranges, oranges to pears, and pears to apples, depending on how the relevant questions are presented to them. The famous materialist, also, as we have seen with the anchoring example, subjective estimates of the number of deaths in Manhattan are influenced by which random number they have just been presented with—the anchor. Given the randomness of the anchor, we will have randomness in the estimates. So if people make inconsistent choices and decisions, the central core of economic optimization fails. You can no longer practice a general theory, and without one you cannot predict. You have to turn to live within a general theory, for Plato's sake.

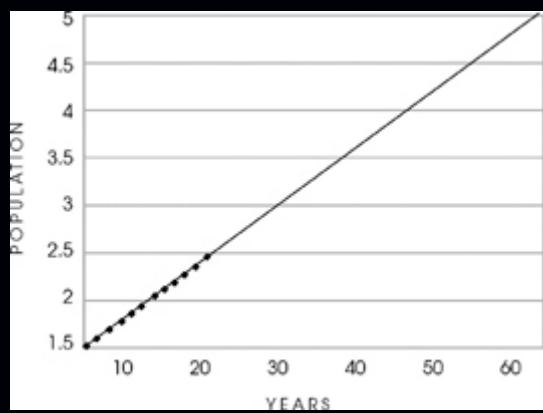
### THE BUSINESS OF EMERALD

Recall the turkey problem. You look at the past and derive some rule about the future. What the problems projecting from the past can be even worse than what we have already learned, because the same past data can confirm a theory and also its exact opposite. If you survive until tomorrow, it could mean that either a) you are more likely to be harvested, b) that you are closer to death. Both conclusions rest on the exact same data. If you are a turkey being fed for a long period of time, you can either naively assume that feeding confirms your safety, or be aware and consider that it confirms the danger of being turned into supper. An acquaintance of mine, a man from India, may indicate his genuine affection for me and his concern for my welfare, if they also confirm his monetary and calculating desire to get my business one day.

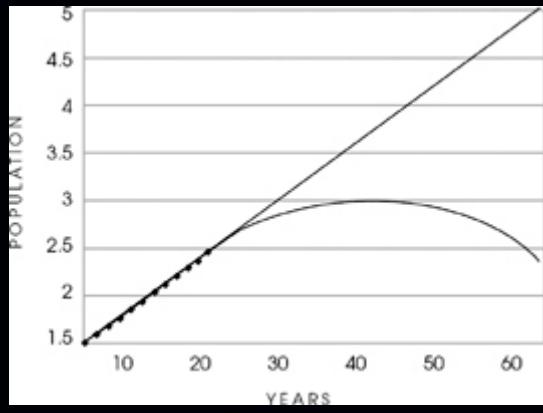
### FIGURE 3



## Chapter 4

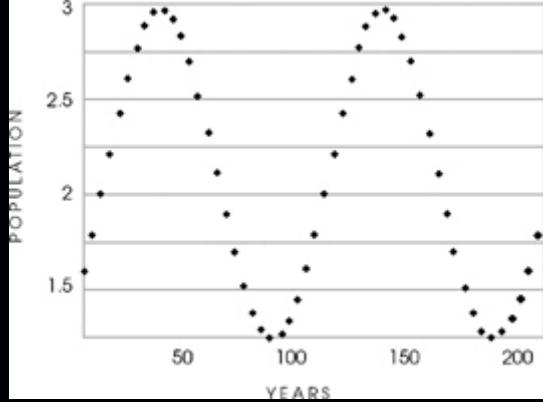


Every day in the world there is one and only one population. But we can project a population into the future. This is what we do in Figure 1.



We look at a broader scale. Now other models also fit the data well.

This is what we do in Figure 2.



And the real generating process is completely unknown. It is nothing to do with a linear model. Some parts of it appear to be linear and we are forced to extrapolating to extremes. \*

So not only can the past be misleading, but there are also many degrees of freedom in our projection of events.

For the technical version of this idea, consider a series of dots on a page representing a number through time—the graph would resemble Figure 1 showing the first thousand days. In Chapter 4 it is said your high school teacher asks you to extrapolate to 1000 days. With a linear model, that is easy. But you can run a straight line, a single straight line from the past to the future. The linear model is just. There is one and only one straight line that can project from a series of points, but it can get broken. If you do not limit yourself to straight lines, you find that there is a huge family of curves that can do the job of connecting the dots. If you project from the past in a linear way, you continue a trend. But possible future deviations from the course of the past are infinite.

This is what the philosopher Nelson Goodman called the riddle of induction. We project a straight line on because we have a linear model in our head—the fact that a number has been for 1000 days straight should make you more confident that it will rise in the future. But if you have a nonlinear model in your head, it might confirm that the number should decline on day 1001.

Let's say that you observe an emerald. It was green yesterday and the day before yesterday. It is green again today. Normally this would confirm the "green" property, we can assume that the emerald will be green tomorrow. But to Goodman, the emerald's color history could equally confirm the "gray" property. What is the gray property? The emerald's gray property is to be green until some specified date, say December 31, 2009.























## THOSE GRAY SWANS OF EXTREMISTAN

It is time to go in some depth into what makes these events occur. First, I have said earlier that the world is moving deeper into Extremistan, but it is less and less governed by Mediocristan—in fact, this idea is more subtle than that. I will show how and present the various ideas we have about the formation of extremes. Second, I have been describing the Gaussian bell curve as a contagious one, several deviations, and it is time to get into that point in some depth. Third, I will present what I call Mandelbrotian, or fractal, randomness. Remember that for an event to be a Black swan, it does not just have to be rare, or just not known to be unexpected, has to be outside our range of possibilities. You must be surprised by it. As it happens, many rare events can yield their structure to us. It is not easy to compute their probability, but it is easy to get a general idea about the possibility of their occurrence. We can turn these Black Swans into Gray Swans, so to speak, reducing their surprise effect. A person aware of the possibility of such events can come to belong to the nonshocker variety.

Finally, I will present the ideas of those philosophers who focus on phony uncertainty. I organized this book in such a way that the more technical (though nonessential) sections are here. They can be skipped without any loss to the thoughtful reader, particularly [Chapters 15, 17](#) and the second half of [Chapter 16](#). I will alert the reader with footnotes. The reader less interested in the mechanics of deviations can then directly proceed to **Part 4**.













Not worth a pinto—Gaussian error. The average man is a monster—but a dirty one—yes or no?—No so different an experiment.

Forget everything you learned in college statistics or probability theory. If you never took such a class, even better, let us start from the very beginning.

### THE GAUSSIAN AND THE MARCHDEUTSCHE

I was walking through the Frankfurt airport in December 2001, on my way from Oslo to Zurich. I had time to kill at the airport and it was a great opportunity for me to buy dark European chocolate, especially since I have managed to successfully convince myself that airport calories don't count. The cashier handed me, among other things, a ten deutschmark bill, an illegal note of which can be seen on the next page. The deutschmark banknotes were going to be put out of circulation in a matter of days, since Europe was moving to the euro. I kept it as a souvenier. Before the arrival of the euro, Europe had plenty of national currencies which was good for printers, money changers, and of course currency traders like this (more or less) human being who was buying my own European chocolate and secretly looking at the bill. I must have been shocked by the fact that there was something written about it. The bill bore the name of Carl Friedrich Gauss and a picture of his Gaussian bell curve.



The last sentence is the most important sentence in this chapter. It is the most important sentence in all of statistics. The striking thing here is that the one possible object that can be linked to the German currency is precisely such a curve. The reichsmark (as the currency was previously called) went from four per dollar to four billion per dollar in the space of a few years during the 1920s, an outcome that tells you that the real cause of the messiness is a description of the randomness in currency fluctuations. All you need to read the bell curve for such a movement to occur once, and only once—just consider the consequences. Yet there was the bell curve, and now to a man Professor Dolan, Geesex, representing a little more than one percent of someone's want to spend time with flogging on a fence, drinking baths, and holding a conversation without a subject. Shockingly, the bell curve is used as a risk-measurement tool by those regulators and central bankers who work with bulls and bears in a boring way about currencies.

### The Increase in the Decrease

The main point of the Gaussian, as I said, is that most observations hover around the mediocre, the average. The odds of a person's decline faster and faster exponential-like as you move away from the average. If you must have only one single piece of information, this is the one: the dramatic increase in the speed of decline the odds as you move away from the center of the average. Look at the bell below for an illustration of this. Let's take an example of a Gaussian quantity, such as height, and applying it a bit to make it more illustrative. Assume that the average height (men and women) is 1.67 meters, or 5 feet 7 inches. Consider what falls a bit off of deviation, here as 10 centimeters, let us look at individuals above 1.67 meters and consider the odds of someone being that tall.\*

10 centimeters taller than the average (i.e., taller than 1.77 m, or 5 feet 9 inches)

20 centimeters taller than the average (i.e., taller than 1.87 m, or 6 feet 1 1/2 inches)

30 centimeters taller than the average (i.e., taller than 1.97 m, or 6 feet 5 1/2 inches)

40 centimeters taller than the average (i.e., taller than 2.07 m, or 6 feet 9 1/2 inches)

50 centimeters taller than the average (i.e., taller than 2.17 m, or 7 feet 1 1/2 inches)

60 centimeters taller than the average (i.e., taller than 2.27 m, or 7 feet 5 1/2 inches)

70 centimeters taller than the average (i.e., taller than 2.37 m, or 7 feet 9 1/2 inches)

80 centimeters taller than the average (i.e., taller than 2.47 m, or 8 feet 1 1/2 inches)

90 centimeters taller than the average (i.e., taller than 2.57 m, or 8 feet 5 1/2 inches)

100 centimeters taller than the average (i.e., taller than 2.67 m, or 8 feet 9 1/2 inches)

100,000,000,000,000,000,000,000

\* and





both knees, garters and ovaries, so that neither should be too sore—unless you get a megajetpack or a megabike on very rare occasion. This would be Mediocristan with a large unit of deviation.

Note once again the following principle: the rarer the event, the higher the error in our estimation of probability—even when using the Gaussian.

Let me show you how the Gaussian bell curve sucks randomness out of life—which is why it is popular. We like it because it allows certainties now! Through averaging, as we discuss next.

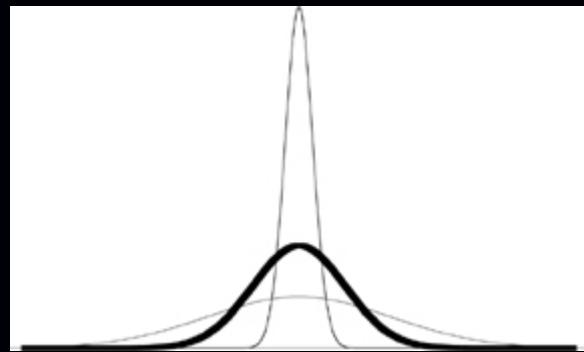
## How Coffee Drinking Can Be Safe

Recall from the Mediocristan discussion in [Chapter 3](#) that no single observation will impact your total. This property will be more and more significant as your samples increase in size. The averages will become more and more stable, to the point where all samples will look alike.

I've had plenty of cups of coffee in my life (as my principal addiction), I have never seen a cup jump two feet from my desk, nor has coffee spilled spontaneously on this manuscript without intervention (even in Russia). Indeed, it will take more than a mild coffee addiction to witness such an event; it would require more caffeine than is perhaps conceivable—the odds are so small, one in so many zeroes, that it would be impossible for me to write them down in my free time.

Yet physical reality makes it possible for my coffee cup to jump—very unlikely, but possible. Particles jump around all the time. How come the coffee cup itself composed of jumping particles, does not? The reason is simply that for the cup to jump would require that all of the particles jump in the same direction and do so in lockstep several times in a row with a compensating move of the table in the opposite direction. All seven billion particles in my coffee cup are not going to jump in the same direction; this is not going to happen in the lifetime of the universe. So I can safely put the coffee cup on the edge of my writing table and worry about more serious sources of uncertainty.

**FIGURE 7: How the Law of Large Numbers Works**



In Mediocristan, as your sample size increases, the distribution becomes more spread out with less and less dispersion—as you can see, the distribution will be narrower and narrower. This, in a nutshell, is how everything in statistical theory works (or is supposed to work). Uncertainty in Mediocristan vanishes under averaging. This illustrates the hallowed “Law of large numbers.”

The safety of my coffee cup illustrates how the randomness of the Gaussian is tamable by averaging. If my cup were one large particle, or even as one, then its jumping would be a problem. But my cup is the sum of billions of very small particles.

Casino operators understand this well, which is why they never if they do things right lose money. They simply do not let one gambler make a massive bet, instead preferring to have plenty of gamblers make series of bets of limited size. Gamblers may bet a total of \$20 million, but you needn't worry about the casino's health; the bets are, say, \$20 on average, the casino keeps the bets as a minimum that will allow the casino owners to sleep right. So the variations in the casino's returns are going to be ridiculously small, no matter the total gambling activity. You will not see anyone leaving the casino with \$1 billion—in the lifetime of the universe.

The above is an application of the statistics of Mediocristan when you have plenty of gamblers; no single gambler will impact the total more than minutely.

The consequence of this is that variations around the average of the Gaussian, also called “errors,” are not the anomalies. They are small and they wash out. They are distributed fluctuations around the mean.

## Love of Certainties

If you ever took a stats class in college, did not understand much of what the professor was excited about, and wondered what “standard deviation” meant, there is nothing to worry about. The notion of standard deviation is meaningless outside of Mediocristan. Clearly it would have been more beneficial and certainly more entertaining, to have taken classes in the neurobiology of aesthetics or postcolonial African dances, and this is easy to see empirically.

Standard deviations do not exist outside the Gaussian, or if they do exist they do not matter and do not explain much. But it gets worse. The Gaussian family (which includes various friends and relatives, such as the Poisson) are the only class of distributions that the standard deviation (and the average) is sufficient to describe. You need nothing else. The bell curve satisfies the reductionism of the decided.

There are other notions that have little or no significance outside of the Gaussian: correlation, and worse,





have exclaimed that if the Greeks had known about it, they would have defined it. His enthusiasm may have contributed to the prevalence of the use of the Gaussian.

Gaukroger was blessed with no mathematical baggage, but he had a rare obsession with measurement. He did not know about the law of large numbers, but rediscovered it from the data itself. He built the quincunx, a pinball machine that shows the development of the bell curve—on which more in a few paragraphs. True, Gaukroger applied the bell curve to areas like genetics and heredity, in which its use was justified. But his enthusiasm helped thrust nascent statistical methods into social issues.

### "Yes/No" Only Please

Let me discuss here the extent of the damage. If you're dealing with qualitative inference, such as in psychology or medicine, looking for yes/no answers to which magnitudes don't apply, then you can assume you're in Medicoverian without serious problems. The impact of the improbable cannot be too large. You have cancer or you don't; you are pregnant or you are not; at work. Degrees of darkness or pregnancy are not relevant unless you are dealing with epidemics. But if you are dealing with aggregates, where magnitudes do matter, such as income, your wealth, return on a portfolio, or book sales, then you will have a problem and get the wrong distribution if you use the Gaussian, as it does not belong there. One single number can disrupt all your averages; one single loss can eradicate a century of profits. You can no longer say this is an exception. The statement "Not a cent less money" is not informative unless you can attach a quantity to that loss. You can lose all your net worth or you can lose a fraction of your daily income; there is a difference.

This explains why empirical psychology and its insights on human nature, which I presented in the earlier parts of this book, are robust to the mistake of using the bell curve; they are also lucky, since most of their variables allow for the application of conventional Gaussian statistics. When measuring how many people in a sample have a disease, or make a mistake, these studies generally elicit a yes/no type of result. A single observation, in itself, can disrupt their overall findings.

I will now proceed to a no-genes presentation of the bell-curve idea from the ground up.

### A LITERARY THOUGHT EXPERIMENT ON HUMAN DATA SETS: CHANCE COMES FROM

Consider a pinball machine like the one shown in Figure 8. Launch 32 balls, assuming a well-balanced board so that the ball has equal odds of falling right or left. You will notice when hitting a pin, your expected outcome is that many balls will land in the center columns, and that the number of balls will decrease as you move to the columns away from the center.

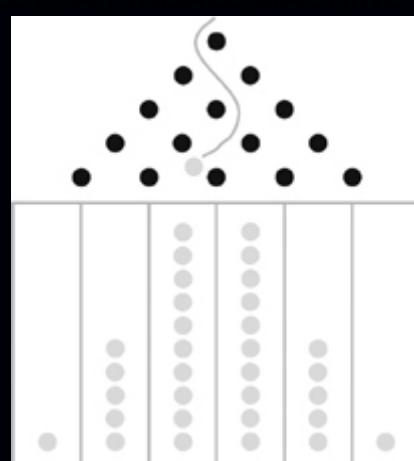
Next, consider a gamblers' thought experiment. A man flips a coin and after each toss he takes a step to the left or a step to the right, depending on whether the coin came up heads or tails. This is called the random walk, but it does not necessarily concern itself with walking. You could identically say that instead of taking a step to the left or to the right, you would win or lose \$1 at every turn, and you will keep track of the cumulative amount that you have in your pocket.

Assume that I set you up in a deadly wager where the odds are neither in your favor nor against you. Flip a coin: heads, you make \$1; tails, you lose \$1.

At the first flip, you will either win or lose.

At the second flip, the number of possible outcomes doubles. Case one: win-win. Case two: win-loss. Case three: loss-win. Case four: loss-loss. Each of these cases has equivalent odds, the combination of a single win and a single loss has an incidence twice as high because cases two and three, win-loss and loss-win, amount to the same outcome. And that is the key for the Gaussian. So much in the middle-washers out—and we will see that there is a lot in the middle, so, if you are playing for \$1 a round, after ten rounds you have a 25 percent chance of making or losing \$0, but a 50 percent chance of breaking even.

FIGURE 8. THE QUINCUNX (SIMPLIFIED)—A PINBALL MACHINE

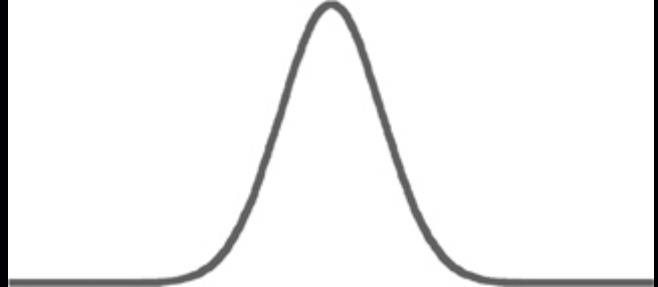


One ball, that is, of every 100 randomly falling in bins, will be the most probable scenario, which greatly resembles the bell curve (aka Gaussian distribution). Courtesy of Alexander Gold.

Let us do another round. The third flip again doubles the number of cases, giving four possible outcomes. Case 1 (W-W-W) wins, win in the second flip branches out into win-win, win-loss and loss-win. We add a win-



Result of forty losses. We see the preo-best curve envelope



Let's keep going. We are 1 in 4,000 times scaling a series of a person's losses above 300 losses as a loss of a penny? As a Platonic form, the pure Gaussian curve is conceptually what happens when we have an infinity of losses per round, with each being infinitesimally small. Do not bother trying to visualize the results, or even make sense out of them. We can no longer talk about an "infinitesimal" last two losses we have an infinity of them, and we are in what mathematicians call a continuous framework. The good news is that there is a substitute. We have moved from a simple bin to something completely smooth. We have moved from observations into the realm of mathematics. In mathematics things have a point to them.

How can you possibly expect to not understand this? Please do not even make an attempt to understand. Just be aware of its use. Think of it as a thermometer; you are not supposed to understand what the temperature means in order to talk about it. You just need to know the correspondence between temperature and comfort for some other empirical consideration. Body degrees correspond to pleasant weather, too below is not something to look forward to. You don't necessarily care about the actual speed of the collisions among particles that more technically explain temperature. Degrees are, in a way, a means for your mind to translate some external phenomena into a number. Likewise, the Gaussian bell curve is set so that 68.2 percent of the observations fall between minus one and plus one standard deviations away from the average. I repeat, do not even try to understand whether standard deviation is average deviation—it is not, and a large (or large) number of people using the word "standard deviation" do not understand this point. Standard deviation is just a number that you scale things to, a matter of mere correspondence if phenomena were Gaussian.

These standard deviations are often nicknamed "sigmas." People also talk about "variance" (some think variance is the square of the sigma, i.e., of the standard deviation).

Note the symmetry in the curve. You get the same results whether the sigma is positive or negative. The odds of falling below -4 sigmas are the same as those of exceeding +4 sigmas, here 1 in 32,000 times.

As the reader can see, the main point of the Gaussian bell curve is, as I have been saying, that most observations hover around the measure, the mean, while the odds of a deviation decline faster and faster (exponentially) as you move away from the mean. If you need to retain one single piece of information, just remember this dramatic speed of decrease in the odds as you move away from the average. Odders are increasingly unlikely. You can safely ignore them.

This property also generates the supreme law of Mediocrity: given the paucity of large deviations, their contribution to the total will be vanishingly small.

In the height example earlier in this chapter, I used units of deviations of ten centimeters, showing how the incidence declined as the height increased. These were one-sigma deviations; the height table also provides an example of the operation of "scaling by a sigma" by using the sigma as a unit of measurement.

### Those Comforting Assumptions

Note the central assumptions we made in the coin-flip game that led to the proto-Gaussian of mid-randomness. First central assumption: the flips are independent of one another. The coin has no memory. The fact that you got heads or tails on the previous flip does not change the odds of your getting heads or tails on the next one. You do not become a better or a bigger over time. If you introduce memory in flips, the entire Gaussian business becomes irrelevant.

Recall our discussions in [Chapter 14](#) on preferential attachment and cumulative advantage. Both theories assert that winning today makes you more likely to win in the future. Therefore, probabilities are dependent on history and the first central assumption leading to the Gaussian bell curve fails in reality. In games, of course, past winnings are not supposed to translate into an increased probability of future gains—but not so in real life, which is why I worry about teaching probability from games. But when winning leads to more winning, you are far more likely to see forty wins in a row than with a proto-Gaussian.

Second central assumption: no "wild" jumps. The step size in the building block of the basic random walk is always known, namely one step. There is no uncertainty as to the size of the step. We did not encounter situations in which the move varied wildly.

Remember that if either of these two central assumptions is not met, your moves (or coin losses) will not cumulatively lead to the bell curve. Depending on what happens, they can lead to the wild Mandelbrotian-style scale-invariant randomness.

### "The Ubiquity of the Gaussian"









Pearls in Science  
What does fractal geometry have to do with the distribution of wealth, the size of cities, returns in the financial markets, the number of casualties in war, or the size of planets? Let us connect the dots.

The key here is that fractal has numerical or statistical measures that are (somewhat) preserved under the scaling transformation unlike the Gaussian. Another view of such self-similarity is presented in Figure 13. We see in Chapter 15 the superpositions are similar to the rich, only richer—wealth is scale invariant, more precisely, of unknown scale dependence.

In the 1960s Mandelbrot presented his ideas on the prices of commodities and financial securities to the economics establishment, and the financial economists got all excited. In 1963 the then-dean of the University of Chicago Graduate School of Business, George Stigler, offered him a professorship. This is the same George Shultz who later became Secretary of State under Ronald Reagan.



FIGURE 13. Apparently, the Gaussian is not the only bell curve. Source: Mandelbrot (1963).

Shultz called him one evening to rescind the offer. At the time of writing, forty-four years later, nothing has happened in economics and social science statistics—except for some cosmetic fiddling that treats the world as if we were subject only to the Gaussian—and yet Nobel medals were being distributed. Some papers were written offering evidence that Mandelbrot was wrong by people who do not get the central argument of this book. You can always prove me wrong by showing that the underlying process is Gaussian by finding periods that do not have rare events, just like you can find an afternoon during which no one killed anyone and use it as “evidence” of how peaceful our world is. I will report that because of the asymmetry with induction, just as it is easier to reject innocence than guilt, it is easier to reject a bell curve than accept it; conversely, it is more difficult to reject a fractal than to accept it. Because a single event can destroy the argument that we face a Gaussian bell curve.

In sum, four decades ago, Mandelbrot gave pearls to economists and resume-building professors, which they rejected because the ideas were too good for them. It was, as the saying goes, “magnetized anti-particles,” and before long,



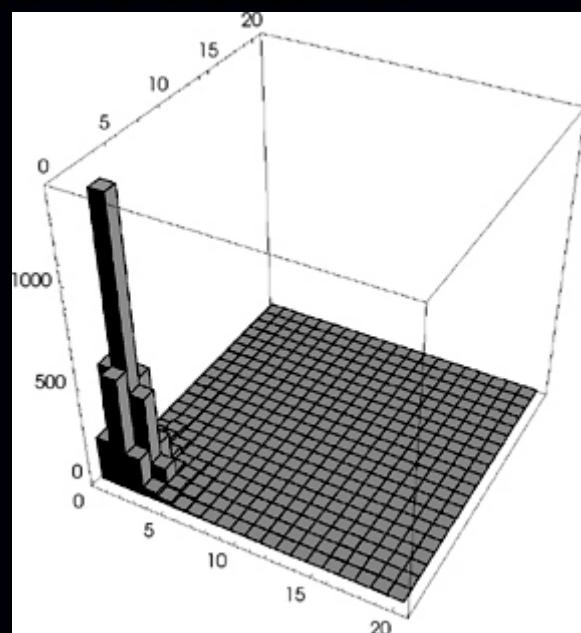
FIGURE 14. The object is not a bell curve. These two photos illustrate what happened to the famous fractal. Courtesy of the author and friends such as Jean-Pierre Hugon, Institut Pluridisciplinaire Hubert Curien, Université de Strasbourg, France.

In the rest of the chapter I will explain how I can endorse Mandelbrot fractals as a representation of randomness without necessarily accepting their precise use. Fractals should be the default, the approach in the framework. They do not solve the Black Swan problem and do not turn off Black Swans and predict events, but they significantly mitigate the Black Swan problem by making such large events considerably rarer than they are. Why? Because only the Swanson give you certainty. More on that later.

## THE LOGIC OF FRACTAL RANDOMNESS WITH A WARNING \*

I have shown in the website ([http://www.johnhorgan.org](#)) Chapter 15 the logic of a fractal distribution. If wealth doubles from 1 million to 2 million, the incidence of people with more than much money is cut in four, which is an exponent of two. If the exponents were one, then the incidence of that wealth or more would be cut in two. The exponent is called the "power" which is why some people use the term "power law". Let us call the number of occurrences higher than a certain level an "exceedance"—an exceedance of two million is the number of persons with wealth more than two million. One more property of these fractals (or another way to express their main property) is that the ratio of two exceedances \* is going to be the ratio of the two numbers to the negative power of the power exponent. Let us illustrate this. Say that you "think" that only 50 books a year will sell more than 200,000 copies, which is what happened last year, and that you "think" that the exponent is around 1.5. You can compute to estimate that around 50 books will sell more than 500,000 copies, or 100,000,000 books. We can continue and note that around 12 books should sell more than a million copies (one 80 times 1,000,000,000 = 80,000,000,000).

FIGURE 15: THE PURE FRACTAL DISTRIBUTION OF WEALTH



The reason for this is that the probability of a person having wealth above a certain level is proportional to the negative power of the wealth level. This is the core of the Black Swan phenomenon. It is the reason that the probability of a person being a millionaire is so small, even though there are many millions of millionaires.

TABLE 2: ASSUMED EXPONENTS FOR VARIOUS PHENOMENA \*

Assumed Exponent (vague approximation)

Population

Books

Books

Books

Books

Books

Books

Books

Books

Books

3 (or lower)

1.5

2 (but possibly a much lower exponent)

\* Source: M.E.J. Newman (2005) and the author's own calculations.

Let me also add that these exponents need not be in terms of numerical precision. We will see why in a minute. But just note at once that we do know *where* these exponents are simply given them, or infer them, for statistical purposes, which makes it hard at times to know the true parameters—if they even exist. Let us first examine the practical consequences of an exponent.

TABLE 3 THE MEANINGS OF THE EXPONENT

Share of the top 1%

99.99%*
99%
47%
34%
27%
22%
10%
6%
4.6%

**Table 3** illustrates the impact of the highly improbable. It shows the contributions of the top 1 percent and 20 percent to the total. The lower the exponent, the higher those contributions. But look how sensitive the process is: between 1.5 and 1.0 you go from 66 percent of the total to 34 percent. Just a 0.2 difference in the exponent changes the result dramatically—and such a difference can come from a simple measurement error. This difference is not mere; just consider that we have no precise idea what the exponent is because we cannot measure it directly. All we do is estimate from past data or rely on theories that allow for the building of some model that would give us some idea—but these models may have hidden weaknesses that prevent us from blindly applying them to reality.

So keep in mind that the 1.5 exponent is an approximation, that it is hard to compute, that you do not get it from the gods, at least not easily, and that you will have a monstrous sampling error. You will observe that the number of books selling above a million copies is not always going to be 7—8 could be as high as 20, or as low as 2.

More significantly, this exponent begins to apply at some number called “crossover,” and addresses numbers larger than this crossover. It may start at 200,000 books, or perhaps only 400,000 books. Likewise, wealth has different properties before, say, \$500 million, when inequality grows, then it does below such a number. How do you know where the crossover point is? This is a problem. My colleagues and I worked with around 20 million pieces of financial data. We all had the same data set, yet we never agreed on exactly what the exponent was in our sets. We knew the data revealed a fractal power law, but we learned that one could not produce a precise number. But what we did know—that the distribution is scaleable and fractal—was sufficient for us to operate and make decisions.

#### The Problem of the Upper Bound

Some people have researched and accepted the fractal “up to a point.” They argue that wealth, book sales, and market returns all have a certain level when things stop being fractal. “Truncation” is what they propose, agreeing that there is a level where fractality might stop, but where? Saying that there is an upper limit, but don’t know how high it is, and saying there is no limit, carry the same consequences in practice. Proposing an upper limit is highly unsafe. You may say, let us cap wealth at \$10 billion in our analyses. Then someone else might say,







What?—Anyone can become president.—Alfred Nobel [1867].—Those medieval days I have in my house two studies, one real, with interesting books and literary material; the other nonliterary, where I do not enjoy working, where I relegated matters, prosaic and narrowly focused. In the nonliterary study is a well full of books on statistics and the history of statistics, books I never had the inclination to burn or throw away, though I find them largely useless outside of their academic applications. (Combes, Chato, and Foucault knew a lot more about probability than all these pseudoscholarly volumes). I cannot use them in class because I promised myself never to teach math, even if dying of starvation. Why can't I use them? Not one of these books deals with Extremistan. Not one. The few books that do are not by mathematicians but by statistical physicists. We are teaching people methods from Mediocristan and turning them loose in Extremistan. It is like developing a medicine for plants and applying it to humans. It is no wonder that we run the biggest risks of all we handle—matters that belong to Extremistan, but treated as if they belonged to Mediocristan, as an "approximation."

Several hundred thousand students in business schools and social science departments from Singapore to Ulsan-Changwon, as well as people in the business world, continue to study "scientific" methods, all grounded in the Gaussian, all embedded in the logic fallacy.

This chapter examines disasters stemming from the application of phony mathematics to social science. The final topic might be the dangers to our society brought about by the Swedish academy that awards the Nobel Prize.

### Only Fifty Years

Let us return to the story of my business life. Look at the graph in [Figure 14](#). In the last fifty years, the ten most extreme days in the financial markets represent half the returns for days in fifty years. Meanwhile, we are trained to think that

Clearly, anyone who wants more than the high number of six sigma as proof that markets are from Extremistan needs to have his head examined. Dozens of papers show the inadequacy of the Gaussian family of distributions and the volatile nature of markets. Recall that, over the years, I myself have run statistical backward and forward on 20 million pieces of data that made me doubt anyone talking about markets in Gaussian terms. But people have a hard time making the leap to the consequences of this knowledge. The strangest thing is that people in business usually agree with me when they listen to me talk or hear me make my case but when they go to the office the next day they revert to the Gaussian tools so entrenched in their heads. Their minds are domain-dependent, so they can exercise critical thinking at a conference while not doing so in the office. Furthermore, the Gaussian tools give them numbers, which seem to be better than nothing. The resulting measure of future uncertainty satisfies our ingrained desire to simplify even if the measure sounding like one single number makes that any less so to be deserved that way.

### The Clinton Reference

I ended [Chapter 1](#) with the stock market crash of 1987, which allowed me to aggressively pursue my Black Swan focus. Since the crash, when I stated that those using sigma (i.e., standard deviations) as a measure of the degree of market randomness were ignoramus, everyone agreed with me. If the world of finance were Gaussian, an episode such as the crash (more than nearly standard deviation) would take place every seven billion lifetimes of the universe (look at the height example in [Chapter 15](#)). According to the circumstances of 1987, people accepted that rare events like price drops are not just a matter of probability. They were also willing to give up on the Gaussian as a central measurement tool—"They've got nothing else." People will stick to another one. Yet the two methods are logically incompatible.

[FIGURE 14](#)



By removing the ten biggest one-day moves from the S&P 500 over the past fifty years, we see a huge difference in what is said vs. what actually happened. Note that the removed moves are more extreme. This is only one of many such tests. When it comes to fitting a Gaussian model, there are many more convincing ones from a mathematical standpoint, such as the incidence of 10-sigma events.)

Unbeknownst to me, 1987 was not the first time the idea of the Gaussian was shown to be faulty. Mandelbrot















Part 4

**THE END**











Postscript Essay

# ON ROBUSTNESS AND FRAGILITY, DEEPER PHILOSOPHICAL AND EMPIRICAL REFLECTIONS



















































First Quadrant	Simple Payoffs	Safe
Second Quadrant	Complex Payoffs	(Sort of) Safe
Third Quadrant	Simple Payoffs	Extremely Safe
Fourth Quadrant	Complex Payoffs	Black Swan Domain
Mediocristan	Mediocristan	Extremistan
Extremistan	Extremistan	Extremistan

**Second Quadrant.** Complex payoffs in Mediocristan statistical methods may work well enough, but there are some risks. The use of Mediocristan models may not be as nuanced, owing to presuppositions about independence, and model error. There clearly are problems here, but these have been addressed extensively in the literature, particularly by David Freedman.

**Third Quadrant.** Simple payoffs in Extremistan, there is little harm in being wrong, because the possibility of extreme events does not impact the payoff. Don't worry too much about Black Swans.

**Fourth Quadrant, the Black Swan Domain.** Complex payoffs in Extremistan. That is where the problem resides: opportunities are present too. We need to avoid prediction of remote payoffs, though not necessarily extreme ones. Payoffs from remote parts of the distribution are more difficult to predict than those from close-by.

\* Interestingly, the author has two parts exposure to positive or negative Black Swans, and focus here on the negative. The positive one is less obvious, and has been discussed in the story of Apollo 13.

## Chapter 13

**TABLE 2: THE FOUR QUADRANTS**

<b>A</b> Mediocristan	I Simple Payoffs	II Complex Payoffs
	First Quadrant <i>Extremely Safe</i>	Second Quadrant <i>(Sort of) Safe</i>
<b>B</b> Extremistan	Third Quadrant <i>Safe</i>	Fourth Quadrant <i>Black Swan Domain</i>















# BEHIND THE CURTAIN: ADDITIONAL NOTES, TECHNICAL COMMENTS, REFERENCES, AND READING RECOMMENDATIONS

I separate topics thematically, so general references will mostly be found in the chapter in which they first occur. I prefer to use a topical sequence here rather than stick to chapter division.

## PROLOGUE and CHAPTER 1

**Bell curve.** When I write “bell curve,” I mean the Gaussian bell curve, a.k.a. normal distribution. All curves look like bells, so this is a nickname. Also, when I write “the Gaussian beam,” I mean all distributions that are similar and for which the improbable is inconsequential and of low impact (more technically, non-scalefree—all moments are finite). Note that the visual presentation of the bell curve in histogram form masks the contribution of the remote events, as such an event will be a point to the far right or far left of the center.

**Diamonds.** See Fox (2002).

**Platonicity.** I’m simply referring to incurring the risk of using a wrong form—not that forms don’t exist, or are not regime-essentials. I am often skeptical of our reverse engineering and identification of the right form. It is an inverse problem!

**Empiricism.** If I call myself an empiricist, or an empirical philosopher, it’s because I am just suspicious of confirmatory generalizations and ready theorizing. Do not confuse this with the British empiricist tradition. Also, many statisticians, as we will see with the Makridakis competition, call themselves “empirical” researchers, but are in fact just the opposite—they fit theories to the past.

**Mention of Christ.** See Flavius Josephus’s *The Jewish War*.

**Great Plan and prediction.** Ferguson (2006b).

**Hindsight bias (retrospective distortion).** See Flaubert (1929).

**Historical fractures.** Braudel (1985), p. 103, quotes a little-known passage from Cauchois. He writes: “This long history wrote since Felix Cauchois, lasted a dozen centuries, longer than the entire history of France.”

**Encountering the past.** And second, the Greek language and thought, all that heritage went up in smoke, as it never happened. (For discussions of discontinuity, see also Gurnevitch (1957), Braudel (1985), Hahn (2004).)

**Religions spread as bestsellers.** Veyne (1974). See also Veyne (2005).

**Clustering in political opinions.** Fisher (2002).

**Categories.** Rosen (1973, 1979). See also Umberto Eco’s *Kant and the Platypus*.

**Historiography and philosophy of history.** Rosen (1973), Carr (1961), Cedars (2002), Braudel (1985a, 1985b), Bourdieu and Martin (1985), Cerboni (1973). MacCannell, Ian Kinsella illustrate the search for causation, which we see already present in Herodotus. For philosophy of history, Rosen (1960a), Fukuyama (1992). For postmodern views, see Jenkins (1993). I show in Part Two how theologographers are unaware of the epistemological difference between forward and backward processes (i.e., between projection and reverse engineering).

**Information and markets.** See Shiller (1981, 1999), Delong et al. (1991), and Culler et al. (1993). The bulk of market moves does not have a reason, just a continued expansion.

**Of descriptive value for crashes.** See Callouts (1997), Shiller (2000), and Kindleberger (2004).

## CHAPTER 2

**Movies.** See De Masi (2002). See also Sargent et al. (2009) for the contagion in movie ratings.

**Religion and domains of cognition.** See Rover (2004).

**Wisdom (madness) of crowds.** Collectively, we can both get wiser or far more foolish. We may collectively have opinions for wisdom-related matters, such as the weight of an ox (see Surprenant, 2004), but my conjecture is that we fail in more complicated predictions (economic variables for which crowds incur pathologies—two heads are worse than one); for decision errors and groups, see Snizker and Buckley (1993). Cf. also Charles Mackay’s *Familiarly Popular Delusions and the Madness of Crowds*.

**Increase in the severity of events.** Zaldenweber (2000).

**Modern life.** The nineteenth-century novelist Emile Zola welcomed the arrival of the market for culture in the late 1800s, of which he seemed to be one of the first beneficiaries. He predicted that the authors and other ability to exploit the commercial system freed them from a dependence on patrons, which, Alas, this was accompanied with more severe concentration—very few people benefited from the system. Laike (2006) shows how most writers, throughout history, have starved. Remarkably, we have ample data from France about the literary tradition.

## CHAPTER 4

**Titanic.** The quote is from Dave Ingram’s presentation at the Enterprise Risk Management Symposium in Chicago on May 2, 2006. For more on ERM, see Lowenstein (2006), Demmer (1999).

**Hume’s exposition.** Hume (1748, 2006).

**Sextus Empiricus.** It is easy, I think, to reject the method of induction. **επαγγελτικός**, or since by way of if they want to make universal convincing on the basis of particulars, they will do this surveying all the particulars or some of them. But if some, the induction will be false, it being that some of the particulars omitted in the induction should be contrary to the universal, and if all, they will labor at an impossible task, since the particulars













occurs in economics, such as tacit knowledge in “know how.” See Ryle (1949), Polanyi (1966, 1974), and Mokyr (2002).

Catherine the Great. The number of lovers comes from Rounding (2008).

Life expectancy. [www.annuityadvantage.com/lifeexpectancy.htm](http://www.annuityadvantage.com/lifeexpectancy.htm) (in projects I have used a probability of exceeding with a certain life expectancy, based on the conditional expectation of it, knowing that it exceeds  $a$ ):

$$E[x|x>a] = \frac{\int_a^\infty xf(x)dx}{\int_a^\infty f(x)dx}.$$

CHAPTERS 11–14

Serenity. See Koessler (1999) and Rees (2004). Rees also has power-law tails on life-expectancy. See also Popper's comments in Popper (2002), and Waller (2002a). Cannon (1940), Mach (1890) (cited in Simonoff, 1999), and Menon and Barber (2004). See Simonoff (2005) for a synthesis. For serenity in medicine and gerontology, see Vaidyanathan (2002).

“Renaissance man.” See [www.bell-labs.com/project/feature/archives/cosmology/](http://www.bell-labs.com/project/feature/archives/cosmology/)

Laser. As usual, there are many stories about the discovery of the laser. In most cases, the actual discoverers are rapidly found, owing to the retrospective distortion. Charles Townsend won the Nobel prize, but was overtaken by his student Gordon Gould, who held that he did the actual work (see “The Economist,” June 8, 2007).

Darwin/Wallace. Quammen (2006).

Popper's attack on historicism. See Popper (2002). Note that I am reinterpreting Popper's idea in a modern manner here, using my own experience and knowledge, not commenting on comments about Popper's work—with the consequent lack of fidelity to his message. In other words, these are not directly Popper's arguments, but largely mine phrased in a Popperian framework. The conditional expectation of an unconditional expectation is an unconditional expectation.

Forecast for the future a hundred years earlier. Bellamy (1888) illustrates our mental projections of the future. However, some stories might be exaggerated. A Patent False Patent Myth will tell a patent official really once resign because he thought nothing was left to invent? Once such myths start they take on a life of their own. *Speciale Inquierer* May–June, 2003.

Observation by Prince. Olson (2006), Prince (1995).

Predicting and explaining. See Thom (1992).

Principles. The three-body problem can be found in Bernoulli (1687), Hooke (1674), and Hooke (1679); Einstein, Poisson (1902). More recent revolutions in Hooke (2004).

Richard Rabb, Berry (1973) and Fagnani and Bonatti (2004).

Very general discussion on complexity. De Bakker (2002), Stenros (1995), and Ruelle (1995). For critics, Barrow (2002).

Hayek. See [www.nobel.se](http://www.nobel.se). Hayek (1945, 1967a) is it that mechanisms do not correct themselves from falling by individual people, but either by the folly of the operators, or something even more severe, by being put out of business? Also, because of contagion, there seems to be little logic to how markets improve, luck plays a part in how soft sciences evolve. See Omerod (2004) for network effects in individuals and societies and the power-law distribution in influence owing to the scale-free aspect of the connections—and the consequential uniqueness. Hayek seems to have been a precursor of Weber's old differentiation between *Naturwissenschaften* and *Geistes Wissenschaften*—but thankfully not Popper.

Insularity of economists. Peters and Baumgartner (2002). One good aspect of the insularity of economists is that they can insult me all they want, with no consequence—it appears that only economists read other economists (so they can write papers for other economists to read). For a more general case, see Matherstein (1999). Note that Groucho taught economic history “It was history.”

Economics as religion. Nelson (2004) and Keen (2004). For methodology, see Blaug (1992). For high priests and lowly philosophers, see Becker, Coase, and Samuelson (2006). Note that the works of Gary Becker and the Platonists of the Chicago School are all marred by the confirmation bias. Becker is quick to show you situations in which people are moved by economic incentives, but does not show you cases (very) more numerous in which people don't care about such materialistic incentives.

The shortest book I've seen in economics is Gove et al. (2006), since it transposes the economic categories in academic economic discourse (one of the authors is the journalist Anatole Kalinowsky). General theory. This fact has not deterred general theorists. One holder of the Platonicizing variety explained to me during a long plane ride from Geneva to New York that the ideas of von Neumann and his colleagues must be rejected because they do not allow us to develop a general equilibrium theory, producing time-inconsistent preferences. For a minute I thought he was joking; he blamed the psychologists' ideas and human incoherence for interfering with the ability to build his Platonic model.

Samuelson. For his optimization, see Samuelson (1937). Also Blaug (1994).

Plato's dogma on body symmetry. Athenian Stranger to Cleonias. In that the right and left hand are supposed





concentration and modern economy. *Res. Economics* (2006).

Choices of society structure and compressed outcomes. The classical paper is Ravinder (1977), though Froncisz, Oppenheimer, and Avery (1987a, 1987b), as well as Lissowski, Tyska, and Chrusz (1993), contradict the notion of the desirability of Rawls's well-thought-by experiment. People prefer maximum average income subjected to a floor constraint or some form of equality for the poor, inequality for the rich type of environment.

Gaussian contagion. Quetelet in Singler (1986). Francis Bacon (as quoted in ten Hockings, "The Tuning of Chance"), "I know of scarcely anything so apt to impress the imagination as the wonderful form of cosmic order expressed by the law of error."

"Finite variance" nonsense. Associated with CLT is an assumption called "finite variance" that is either technical; none of these building block steps can take an infinite value if you square them or multiply them by themselves. They need to be bounded at some number. We simplified here by making them all one single step of finite standard deviation. But the problem is that some fractal powerlaws may have finite variance, but still not take us there rapidly. See Bouchaud and Potters (2000).

Lognormal. There is an intermediate variety that is called the lognormal, emphasized by one Gilbreath (see Sauer (1987)), early in the twentieth century as an attempt to explain the distribution of wealth. In this framework, it is not quite that the wealthy get wealthier, in a pure preferential attachment situation, but that if your wealth is at 100 you will vary by 1, but when your wealth is at 1,000, you will vary by 10. The relative changes in your wealth are Gaussian. So the lognormal superficially resembles the fractal, in the sense that it may tolerate some large deviations, but it is dangerous because these rapidly taper off in the end. The introduction of the lognormal was a very bad compromise, but a way to conceal the flaws of the Gaussian.

Extinctions. Stanley (2004), for extinctions from abrupt fractures, see Courtillot (1996) and Courtillot and Gaillard (1996); James, Fidinger, and Gould.

## FRACTALS, POWER LAWS, and SCALE-FREE DISTRIBUTIONS

Definition. Technically,  $P(x) \propto x^{-\alpha}$ , where  $\alpha$  is supposed to be the power-law exponent. It is said to be scale-free, in the sense that it does not have a characteristic scale: relative deviation  $\frac{P_{>x}}{P_{>nx}}$  does not depend on  $x$ , but on  $n$ —for  $x$  "large enough." Now, in the other class of distribution, the one that can intuitively describe as conceivable, with the typical shape  $p(x) = \exp(-x)$ , the scale will be  $x$ .

Problem of "how large": how the problem that is usually misunderstood. This scalability might stop somewhere, but I do not know where, so I might consider it infinite. The statements "very large" and "I don't know how large" and "infinitely large" are epistemologically substitutable. There might be a point at which the distribution is no longer scale-free; we look at them more graphically.

Log-log plots. If  $P(x) \propto x^{-\alpha}$  is available, when we do a log-log plot (i.e., plot  $\ln P(x)$  and  $\ln x$  on a logarithmic scale), see Figures 15 and 16, we should see a straight line.

Fractals and power laws. Mandelbrot (1975, 1982), Schrodinger (1951) is imperative. John Chapman's unpublished manuscript "The Fractal Heritage" (Chapman (2006)) is the best review piece I've seen. See also Mitznerwacher (2003).

"To come very near true theory and to grasp its precise application are two very different things, as the history of science teaches us. Everything of importance has been said before, by somebody who did not discover it." Whitehead (1926).

Fractals in poetry. For the quote on Dickinson, see Fillion (1998).

Lacunarity. Broekman (2006), in the arts, Mandelbrot (1982).

Fractals in medicine. "New Tools to Diagnose and Treat Breast Cancer," *Newswise*, July 10, 2006.

General reference books in statistical physics. The most complete in relation to fat tails is Bouchet (2004). See also Van (2001) or the far deeper Bouchaud and Potters (2002) in financial prices and econophysics. For "complexity" theory, technical books: Boccaletti (2004), Shroeder (1994), the popular Huillet (1994), and also Prigogine (1996).

Fitting processes. For the philosophy of the problem, Taleb and Piper (2004). See also Pirogovskiy and Somoto (2004), Somoto et al. (2004), and Somoto and Iba (2001).

Poisson jumps. Sometimes people propose a Gaussian distribution with a small probability of a "Poisson" jump. This may be fine, but how do you know how large the jump is going to be? Past data might not tell you how large the jumps are.

Small sample effects. Wilson (2004), Officer (1972) is quite ignorant of the point.

Recursivity of statistics. Taleb and Piper (2004), Glynn et al. (2005).

Biology. Modern molecular biology pioneers Salvador Luria and Max Delbrück witnessed a clustering phenomenon with the occasional occurrence of extremely large mutants in a bacterial colony, larger than all other bacteria.

Thermodynamics. Energy maximization without the constraint of a second moment. The two exclusive domains of abstraction, vertical or straight line with slopes either negative infinity or constant negative  $\alpha$ . Note that since probabilities need to add up to 1 (even in France) there cannot be other alternatives to the two beams, which is why I narrow it down to these two exclusively.

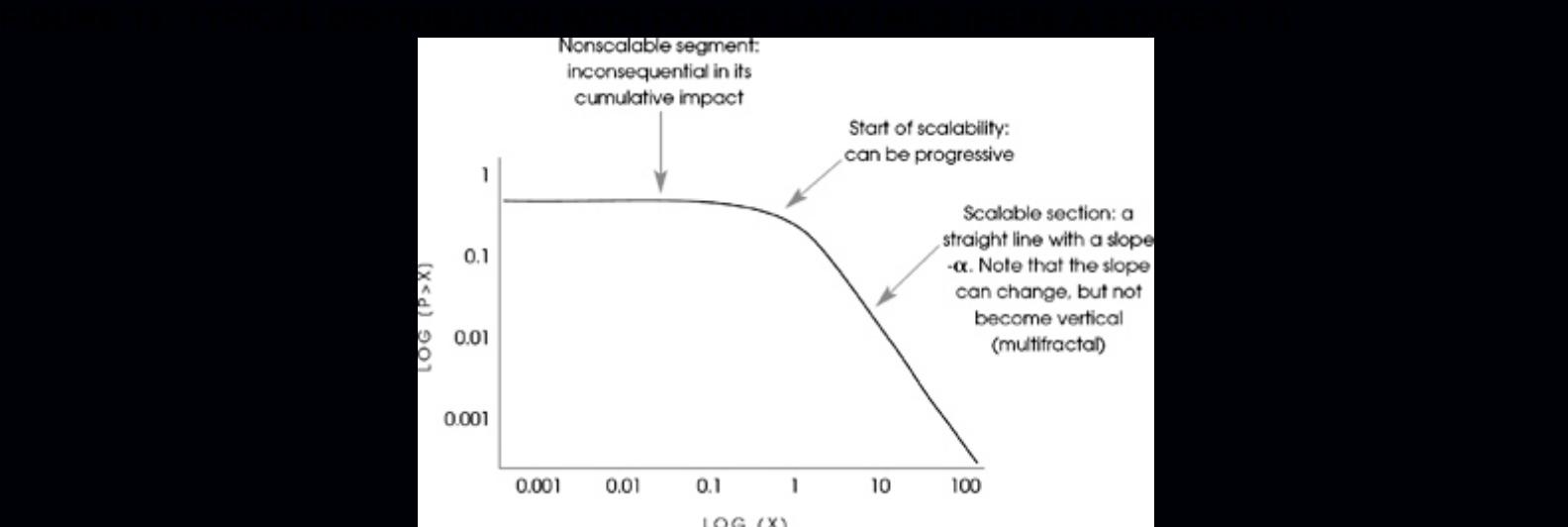


FIGURE 16

There are very simple and very complex powerlaws—according to the process of self-organized criticality, as in using to the scarcity of data on the far right, leads to a very different Mandelbrot's theory of 1952 (see Mandelbrot, 1957a). Below a more sophisticated view of events by Gellman.

**Imitation chains and pathologies.** An informational cascade is a process where a purely rational agent makes a particular choice ignoring his own private information or judgment in favor of others. You can follow your friends, you may be aware of a danger, may be missing. It is efficient to do what others do instead of having to reinvent the wheel every time. But this copying the behavior of others can lead to imitation chains, when everyone is moving in the same direction, and it can be for spurious reasons. This behavior causes some market bubbles and the formation of massive cultural fads. Bakhtin et al. (1992), in psychology see Hansen and Deci (1978), in biology selection, Dugatkin (2003), Kirkpatrick and Dugatkin (1994), Self-organized criticality, Bak and Chen (1994), Bak (1996).

**Economic variables.** Gundlach and Murphy (2003). Most economic variables seem to follow a "stable" distribution. They include foreign exchange, the GDP, the money supply, interest rates (long and short term), and industrial production.

Statistics not accepting scalability. Flawed reasoning resulting for sampling error in the tail and boundedness. Perini (2000), for instance, does not understand the difference between absence of evidence and evidence of absence.

**Time series and memory.** You can have fractal memory, i.e., the effect of past events on the present has an impact that goes "all the way down" as power-law, not exponential.

Marmott's work, Marmott (2004).

#### REFERENCES

- Economists, Weintraub (2002), Sennholz (1992).
- Portfolio theory and modern finance, Markowitz (1952, 1959), Huizinga and Litzenberger (1985) and Sharpe (1994, 1999). What is called the Sharpe ratio is meaningless outside of Markowitz. The contents of Steve Ross's book Ross (2004) on "neoclassical finance" are completely canceled if you consider extremism in spite of the "elegant" mathematics and the beautiful top-down theories. Anecdote of Nelson in Merton (1992).
- Obsession with measurement, Crosby (1987) is often shown to me as convincing evidence that measuring was a great accomplishment not knowing that it applies to Medicine and Medicine only. Bernstein (1996) makes the same error.
- Power laws in finance, Mandelbrot (1963), Gabaix et al. (2004), and Stanley et al. (2000). Kondor and Kondor (2004), Vinken and Vinken (2002), Landwehr, Matalas (2003), Regnault, Souza and Polson (2003).























































PHOTO: © SARAH JOSEPHINE TALEB

NASSIM NICHOLAS TALEB has devoted his life to immersing himself in problems of luck, uncertainty, probability, and knowledge, and he has led three high-profile careers around his ideas: as a man of letters, as a business raconteur, and as a university professor. Although he spends most of his time as a hermit, meditating in cafes across the planet, he is currently Distinguished Professor at New York University's Polytechnic Institute and a principal of Universe LP. His books, *Fat tail vs. Thin tails* and *The Black Swan*, have been published in thirty-one languages.

