Fooled By Randomness

Nine

•

IT IS EASIER TO BUY AND SELL THAN FRY AN EGG

Some technical extensions of the survivorship bias. On the distribution of “coincidences” in life. It is preferable to be lucky than competent (but you can be caught). The birthday paradox. More charlatans (and more journalists). How the researcher with work ethics can find just about anything in data. On dogs not barking.

This afternoon I have an appointment with my dentist (it will mostly consist in the dentist picking my brain on Brazilian bonds). I can state with a certain level of comfort that he knows something about teeth, particularly if I enter his office with a toothache and exit it with some form of relief. It will be difficult for someone who knows literally nothing about teeth to provide me with such relief, except if he is particularly lucky on that day—or has been very lucky in his life to become a dentist while not knowing anything about teeth. Looking at his diploma on the wall, I determine that the odds that he repeatedly gave correct answers to the exam questions and performed satisfactorily on a few thousand cavities before his graduation—out of plain randomness—are remarkably small.

Later in the evening, I go to Carnegie Hall. I can say little about the pianist; I even forgot her unfamiliar foreign-sounding name. All I know is that she studied in some Muscovite conservatory. But I can expect to get some music out of the piano. It will be rare to have someone who performed brilliantly enough in the past to get to Carnegie Hall and now turns out to have benefited from luck alone. The expectation of having a fraud who will bang on the piano, producing only cacophonous sounds, is indeed low enough for me to rule it out completely.

I was in London last Saturday. Saturdays in London are magical; bustling but without the mechanical industry of a weekday or the sad resignation of a Sunday. Without a wristwatch or a plan I found myself in front of my favorite carvings by Canova at the Victoria and Albert Museum. My professional bent immediately made me question whether randomness played a large role in the carving of these marble statues. The bodies were realistic reproductions of human figures, except that they were more harmonious and finely balanced than anything I have seen mother nature produce on its own (Ovid’s materiam superabat opus comes to mind). Could such finesse be a product of luck?

I can practically make the same statement about anyone operating in the physical world, or in a business in which the degree of randomness is low. But there is a problem in anything related to the business world. I am bothered because tomorrow, unfortunately, I have an appointment with a fund manager seeking my help, and that of my friends, in finding investors. He has what he claims is a good track record. All I can infer is that he has learned to buy and sell. And it is harder to fry an egg than buy and sell. Well . . . the fact that he made money in the past may have some relevance, but not terribly so. This is not to say that it is always the case; there are some instances in which one can trust a track record, but, alas, there are not too many of these. As the reader now knows, the fund manager can expect to be heckled by me during the presentation, particularly if he does not exhibit the minimum of humility and self-doubt that I would expect from someone practicing randomness. I will probably bombard him with questions that he may not be prepared to answer, blinded by his past results. I will probably lecture him that Machiavelli ascribed to luck at least a 50% role in life (the rest was cunning and bravura), and that was before the creation of modern markets.

In this chapter, I discuss some well-known counterintuitive properties of performance records and historical time series. The concept presented here is well-known for some of its variations under the

names survivorship bias, data mining, data snooping, over-fitting, regression to the mean, etc., basically situations where the performance is exaggerated by the observer, owing to a misperception of the importance of randomness. Clearly, this concept has rather unsettling implications. It extends to more general situations where randomness may play a share, such as the choice of a medical treatment or the interpretation of coincidental events.

When I am tempted to suggest a possible future contribution of financial research to science in general, I adduce the analysis of data mining and the study of survivorship biases. These have been refined in finance but can extend to all areas of scientific investigation. Why is finance so rich a field? Because it is one of the rare areas of investigation where we have plenty of information (in the form of abundant price series), but no ability to conduct experiments as in, say, physics. This dependence on past data brings about its salient defects.

FOOLED BY NUMBERS

Placebo Investors

I have often been faced with questions of the sort: “Who do you think you are to tell me that I might have been plain lucky in my life?” Well, nobody really believes that he or she was lucky. My approach is that, with our Monte Carlo engine, we can manufacture purely random situations. We can do the exact opposite of conventional methods; in place of analyzing real people hunting for attributes we can create artificial ones with precisely known attributes. Thus we can manufacture situations that depend on pure, unadulterated luck, without the shadow of skills or whatever we have called nonluck in Table P.1. In other words, we can man-make pure nobodies to laugh at; they will be by design stripped of any shadow of ability (exactly like a placebo drug).

We saw in Chapter 5 how people may survive owing to traits that momentarily fit the given structure of randomness. Here we take a far simpler situation where we know the structure of randomness; the first such exercise is a finessing of the old popular saying that even a broken clock is right twice a day. We will take it a bit further to show that statistics is a knife that cuts on both sides. Let us use the Monte Carlo generator introduced earlier and construct a population of 10,000 fictional investment managers (the generator is not terribly necessary since we can use a coin, or even do plain algebra, but it is considerably more illustrative—and fun). Assume that they each have a perfectly fair game; each one has a 50% probability of making $10,000 at the end of the year, and a 50% probability of losing $10,000. Let us introduce an additional restriction; once a manager has a single bad year, he is thrown out of the sample, good-bye and have a nice life. Thus we will operate like the legendary speculator George Soros who was said to tell his managers gathered in a room: “Half of you guys will be out by next year” (with an Eastern European accent). Like Soros, we have extremely high standards; we are looking only for managers with an unblemished record. We have no patience for low performers.

The Monte Carlo generator will toss a coin; heads and the manager will make $10,000 over the year, tails and he will lose $10,000. We run it for the first year. At the end of the year, we expect 5,000 managers to be up $10,000 each, and 5,000 to be down $10,000. Now we run the game a second year. Again, we can expect 2,500 managers to be up two years in a row; another year, 1,250; a fourth one, 625; a fifth, 313. We have now, simply in a fair game, 313 managers who made money for five years in a row. Out of pure luck.

Meanwhile if we throw one of these successful traders into the real world we would get very interesting and helpful comments on his remarkable style, his incisive mind, and the influences that helped him achieve such success. Some analysts may attribute his achievement to precise elements among his childhood experiences. His biographer will dwell on the wonderful role models provided by his parents; we would be supplied with black-and-white pictures in the middle of the book of a great mind in the making. And the following year, should he stop outperforming (recall that his odds of having a good

year have stayed at 50%) they would start laying blame, finding fault with the relaxation in his work ethics, or his dissipated lifestyle. They will find something he did before when he was successful that he has subsequently stopped doing, and attribute his failure to that. The truth will be, however, that he simply ran out of luck.

Nobody Has to Be Competent

Let's push the argument further to make it more interesting. We create a cohort that is composed exclusively of incompetent managers. We will define an incompetent manager as someone who has a negative expected return, the equivalent of the odds being stacked against him. We instruct the Monte Carlo generator now to draw from an urn. The urn has 100 balls, 45 black and 55 red. By drawing with replacement, the ratio of red to black balls will remain the same. If we draw a black ball, the manager will earn $10,000. If we draw a red ball, he will lose $10,000. The manager is thus expected to earn $10,000 with 45% probability, and lose $10,000 with 55%. On average, the manager will lose $1,000 each round—but only on average.

At the end of the first year, we still expect to have 4,500 managers turning a profit (45% of them), the second, 45% of that number, 2,025. The third, 911; the fourth, 410; the fifth, 184. Let us give the surviving managers names and dress them in business suits. True, they represent less than 2% of the original cohort. But they will get attention. Nobody will mention the other 98%.What can we conclude?

The first counterintuitive point is that a population entirely composed of bad managers will produce a small amount of great track records. As a matter of fact, assuming the manager shows up unsolicited at your door, it will be practically impossible to figure out whether he is good or bad. The results would not markedly change even if the population were composed entirely of managers who are expected in the long run to lose money. Why? Because owing to volatility, some of them will make money. We can see here that volatility actually helps bad investment decisions.

The second counterintuitive point is that the expectation of the maximum of track records, with which we are concerned, depends more on the size of the initial sample than on the individual odds per manager. In other words, the number of managers with great track records in a given market depends far more on the number of people who started in the investment business (in place of going to dental school), rather than on their ability to produce profits. It also depends on the volatility. Why do I use the notion of expectation of the maximum? Because I am not concerned at all with the average track record. I will get to see only the best of the managers, not all of the managers. This means that we would see more "excellent managers" in 2006 than in 1998, provided the cohort of beginners was greater in 2001 than it was in 1993—I can safely say that it was.

Regression to the Mean

The "hot hand in basketball" is another example of misperception of random sequences: It is very likely in a large sample of players for one of them to have an inordinately lengthy lucky streak. As a matter of fact it is very unlikely that an unspecified player somewhere does not have an inordinately lengthy lucky streak. This is a manifestation of the mechanism called regression to the mean. I can explain it as follows:

Generate a long series of coin flips producing heads and tails with 50% odds each and fill up sheets of paper. If the series is long enough you may get eight heads or eight tails in a row, perhaps even ten of each. Yet you know that in spite of these wins the conditional odds of getting a head or a tail is still 50%. Imagine these heads and tails as monetary bets filling up the coffers of an individual. The deviation from the norm as seen in excess heads or excess tails is here entirely attributable to luck, in other words, to variance, not to the skills of the hypothetical player (since there is an even probability of getting either).

A result is that in real life, the larger the deviation from the norm, the larger the probability of it coming from luck rather than skills: Consider that even if one has 55% probability of heads, the odds of ten wins is still very small. This can be easily verified in stories of very prominent people in trading rapidly reverting to obscurity, like the heroes I used to watch in trading rooms. This applies to height of individuals or the size of dogs. In the latter case, consider that two average-sized parents produce a large litter. The largest dogs, if they diverge too much from the average, will tend to produce offspring of smaller size than themselves, and vice versa. This “reversion” for the large outliers is what has been observed in history and explained as regression to the mean. Note that the larger the deviation, the more important its effect.

Again, one word of warning: All deviations do not come from this effect, but a disproportionately large proportion of them do.

Ergodicity

To get more technical, I have to say that people believe that they can figure out the properties of the distribution from the sample they are witnessing. When it comes to matters that depend on the maximum, it is altogether another distribution that is being inferred, that of the best performers. We call the difference between the average of such distribution and the unconditional distribution of winners and losers the survivorship bias—here the fact that about 3% of the initial cohort discussed earlier will make money five years in a row. In addition, this example illustrates the properties of ergodicity, namely, that time will eliminate the annoying effects of randomness. Looking forward, in spite of the fact that these managers were profitable in the past five years, we expect them to break even in any future time period. They will fare no better than those of the initial cohort who failed earlier in the exercise. Ah, the long term.

A few years ago, when I told one A., a then Master-of-the-Universe type, that track records were less relevant than he thought, he found the remark so offensive that he violently flung his cigarette lighter in my direction. The episode taught me a lot. Remember that nobody accepts randomness in his own success, only his failure. His ego was pumped up as he was heading up a department of “great traders” who were then temporarily making a fortune in the markets and attributing the idea to the soundness of their business, their insights, or their intelligence. They subsequently blew up during the harsh New York winter of 1994 (it was the bond market crash that followed the surprise interest rate hike by Alan Greenspan). The interesting part is that several years later I can hardly find any of them still trading (ergodicity).

Recall that the survivorship bias depends on the size of the initial population. The information that a person derived some profits in the past, just by itself, is neither meaningful nor relevant. We need to know the size of the population from which he came. In other words, without knowing how many managers out there have tried and failed, we will not be able to assess the validity of the track record. If the initial population includes ten managers, then I would give the performer half my savings without a blink. If the initial population is composed of 10,000 managers, I would ignore the results. The latter situation is generally the case; these days so many people have been drawn to the financial markets. Many college graduates are trading as a first career, failing, then going to dental school.

If, as in a fairy tale, these fictional managers materialized into real human beings, one of these could be the person I am meeting tomorrow at 11:45 a.m. Why did I select 11:45 a.m.? Because I will question him about his trading style. I need to know how he trades. I will then be able to claim that I have to rush to a lunch appointment if the manager puts too much emphasis on his track record.

LIFE IS COINCIDENTAL

Next, we look at the extensions to real life of our bias in the understanding of the distribution of coincidences.

The Mysterious Letter

You get an anonymous letter on January 2 informing you that the market will go up during the month. It proves to be true, but you disregard it owing to the well-known January effect (stocks have gone up historically during January). Then you receive another one on February 1 telling you that the market will go down. Again, it proves to be true. Then you get another letter on March 1—same story. By July you are intrigued by the prescience of the anonymous person and you are asked to invest in a special offshore fund. You pour all your savings into it. Two months later, your money is gone. You go spill your tears on your neighbor's shoulder and he tells you that he remembers that he received two such mysterious letters. But the mailings stopped at the second letter. He recalls that the first one was correct in its prediction, the other incorrect.

What happened? The trick is as follows. The con operator pulls 10,000 names out of a phone book. He mails a bullish letter to one half of the sample, and a bearish one to the other half. The following month he selects the names of the persons to whom he mailed the letter whose prediction turned out to be right, that is, 5,000 names. The next month he does the same with the remaining 2,500 names, until the list narrows down to 500 people. Of these there will be 200 victims. An investment in a few thousand dollars' worth of postage stamps will turn into several million.

An Interrupted Tennis Game

It is not uncommon for someone watching a tennis game on television to be bombarded by advertisements for funds that did (until that minute) outperform others by some percentage over some period. But, again, why would anybody advertise if he didn't happen to outperform the market? There is a high probability of the investment coming to you if its success is caused entirely by randomness. This phenomenon is what economists and insurance people call adverse selection. Judging an investment that comes to you requires more stringent standards than judging an investment you seek, owing to such selection bias. For example, by going to a cohort composed of 10,000 managers, I have 2/100 chances of finding a spurious survivor. By staying home and answering my doorbell, the chance of the soliciting party being a spurious survivor is closer to 100%.

Reverse Survivors

We have so far discussed the spurious survivor—the same logic applies to the skilled person who has the odds markedly stacked in her favor, but who still ends up going to the cemetery. This effect is the exact opposite of the survivorship bias. Consider that all one needs is two bad years in the investment industry to terminate a risk-taking career and that, even with great odds in one's favor, such an outcome is very possible. What do people do to survive? They maximize their odds of staying in the game by taking black-swan risks (like John and Carlos)—those that fare well most of the time, but incur a risk of blowing up.

The Birthday Paradox

The most intuitive way to describe the data mining problem to a nonstatistician is through what is called the birthday paradox, though it is not really a paradox, simply a perceptional oddity. If you meet someone randomly, there is a one in 365.25 chance of your sharing their birthday, and a considerably smaller one of having the exact birthday of the same year. So, sharing the same birthday would be a coincidental event that you would discuss at the dinner table. Now let us look at a situation where there are 23 people in a room. What is the chance of there being 2 people with the same birthday? About 50%. For we are not specifying which people need to share a birthday; any pair works.

It's a Small World!

A similar misconception of probabilities arises from the random encounters one may have with relatives or friends in highly unexpected places. "It's a small world!" is often uttered with surprise. But these are not improbable occurrences—the world is much larger than we think. It is just that we are not truly testing for the odds of having an encounter with one specific person, in a specific location at a specific time. Rather, we are simply testing for any encounter, with any person we have ever met in the past, and in any place we will visit during the period concerned. The probability of the latter is considerably higher, perhaps several thousand times the magnitude of the former.

When the statistician looks at the data to test a given relationship, say, to ferret out the correlation between the occurrence of a given event, like a political announcement, and stock market volatility, odds are that the results can be taken seriously. But when one throws the computer at data, looking for just about any relationship, it is certain that a spurious connection will emerge, such as the fate of the stock market being linked to the length of women's skirts. And just like the birthday coincidences, it will amaze people.

Data Mining, Statistics, and Charlatanism

What is your probability of winning the New Jersey lottery twice? One in 17 trillion. Yet it happened to Evelyn Adams, whom the reader might guess should feel particularly chosen by destiny. Using the method we developed above, researchers Percy Diaconis and Frederick Mosteller estimated at 30 to 1 the probability that someone, somewhere, in a totally unspecified way, gets so lucky!

Some people carry their data mining activities into theology—after all, ancient Mediterraneans used to read potent messages in the entrails of birds. An interesting extension of data mining into biblical exegesis is provided in The Bible Code by Michael Drosnin. Drosnin, a former journalist (seemingly innocent of any training in statistics), aided by the works of a "mathematician," helped "predict" the former Israeli Prime Minister Yitzhak Rabin's assassination by deciphering a bible code. He informed Rabin, who obviously did not take it too seriously. The Bible Code finds statistical irregularities in the Bible; these help predict some such events. Needless to say that the book sold well enough to warrant a sequel predicting with hindsight even more such events.

The same mechanism is behind the formation of conspiracy theories. Like The Bible Code they can seem perfect in their logic and can cause otherwise intelligent people to fall for them. I can create a conspiracy theory by downloading hundreds of paintings from an artist or group of artists and finding a constant among all those paintings (among the hundreds of thousand of traits). I would then concoct a conspiratorial theory around a secret message shared by these paintings. This is seemingly what the author of the bestselling The Da Vinci Code did.

The Best Book I Have Ever Read!

My favorite time is spent in bookstores, where I aimlessly move from book to book in an attempt to make a decision as to whether to invest the time in reading it. My buying is frequently made on impulse, based on superficial but suggestive clues. Frequently, I have nothing but a book jacket as appendage to my decision making. Jackets often contain praise by someone, famous or not, or excerpts from a book review. Good praise by a famous and respected person or a well-known magazine would sway me into buying the book.

What is the problem? I tend to confuse a book review, which is supposed to be an assessment of the quality of the book, with the best book reviews, marred with the same survivorship biases. I mistake the distribution of the maximum of a variable with that of the variable itself. The publisher will never put on

the jacket of the book anything but the best praise. Some authors go even a step beyond, taking a tepid or even unfavorable book review and selecting words in it that appear to praise the book. One such example came from one Paul Wilmott (an English financial mathematician of rare brilliance and irreverence) who managed to announce that I gave him his “first bad review,” yet used excerpts from it as praise on the book jacket (we later became friends, which allowed me to extract an endorsement from him for this book).

The first time I was fooled by this bias was upon buying, when I was sixteen, Manhattan Transfer, a book by John Dos Passos, the American writer, based on praise on the jacket by the French writer and “philosopher” Jean-Paul Sartre, who claimed something to the effect that Dos Passos was the greatest writer of our time. This simple remark, possibly blurted out in a state of intoxication or extreme enthusiasm, caused Dos Passos to become required reading in European intellectual circles, as Sartre’s remark was mistaken for a consensus estimate of the quality of Dos Passos rather than what it was, the best remark. (In spite of such interest in his work, Dos Passos has reverted to obscurity.)

The Backtester

A programmer helped me build a backtester. It is a software program connected to a database of historical prices, which allows me to check the hypothetical past performance of any trading rule of average complexity. I can just apply a mechanical trading rule, like buy NASDAQ stocks if they close more than 1.83% above their average of the previous week, and immediately get an idea of its past performance. The screen will flash my hypothetical track record associated with the trading rule. If I do not like the results, I can change the percentage to, say, 1.2%. I can also make the rule more complex. I will keep trying until I find something that works well.

What am I doing? The exact same task of looking for the survivor within the set of rules that can possibly work. I am fitting the rule on the data. This activity is called data snooping. The more I try, the more I am likely, by mere luck, to find a rule that worked on past data. A random series will always present some detectable pattern. I am convinced that there exists a tradable security in the Western world that would be 100% correlated with the changes in temperature in Ulan Bator, Mongolia.

To get technical, there are even worse extensions. An outstanding recent paper by Sullivan, Timmerman, and White goes further and considers that the rules that may be in use successfully today may be the result of a survivorship bias.

Suppose that, over time, investors have experimented with technical trading rules drawn from a very wide universe—in principle thousands of parameterizations of a variety of types of rules. As time progresses, the rules that happen to perform well historically receive more attention and are considered “serious contenders” by the investment community, while unsuccessful trading rules are more likely to be forgotten. . . . If enough trading rules are considered over time, some rules are bound by pure luck, even in a very large sample, to produce superior performance even if they do not genuinely possess predictive power over asset returns. Of course, inference based solely on the subset of surviving trading rules may be misleading in this context since it does not account for the full set of initial trading rules, most of which are unlikely to have underperformed.

I have to decry some excesses in backtesting that I have closely witnessed in my private career. There is an excellent product designed just for that, called Omega TradeStation, that is currently on the market, in use by tens of thousands of traders. It even offers its own computer language. Beset with insomnia, the computderized day traders become night testers plowing the data for some of its properties. By dint of throwing their monkeys on typewriters, without specifying what book they want their monkey to write, they will hit upon hypothetical gold somewhere. Many of them blindly believe in it.

One of my colleagues, a man with prestigious degrees, grew to believe in such a virtual world to the point of losing all sense of reality. Whether the modicum of common sense left in him might have rapidly vanished under the mounds of simulations, or whether he might have had none to engage in such pursuit, I cannot tell. By closely watching him I learned that what natural skepticism he may have had vanished under the weight of data—for he was extremely skeptical, but in the wrong area. Ah, Hume!

A More Unsettling Extension

Historically, medicine has operated by trial and error—in other words, statistically. We know by now that there can be entirely fortuitous connections between symptoms and treatment, and that some medications succeed in medical trials for mere random reasons. I cannot claim expertise in medicine, but have been a steady reader of a segment of the medical literature over the past half decade, long enough to be concerned with the standards, as we will see in the next chapter. Medical researchers are rarely statisticians; statisticians are rarely medical researchers. Many medical researchers are not even remotely aware of this data mining bias. True, it may only play a small role, but it is certainly present. One recent medical study links cigarette smoking to a reduction in breast cancer, thus conflicting with all previous studies. Logic would indicate that the result may be suspicious, the result of mere coincidence.

The Earnings Season: Fooled by the Results

Wall Street analysts, in general, are trained to find the accounting tricks that companies use to hide their earnings. They tend to (occasionally) beat the companies at that game. But they are neither trained to reflect nor to deal with randomness (nor to understand the limitations of their methods by introspecting—stock analysts have both a worse record and higher idea of their past performance than weather forecasters). When a company shows an increase in earnings once, it draws no immediate attention. Twice, and the name starts showing up on computer screens. Three times, and the company will merit some buy recommendation.

Just as with the track record problem, consider a cohort of 10,000 companies that are assumed on average to barely return the risk-free rate (i.e., Treasury bonds). They engage in all forms of volatile business. At the end of the first year, we will have 5,000 “star” companies showing an increase in profits (assuming no inflation), and 5,000 “dogs.” After three years, we will have 1,250 “stars.” The stock review committee at the investment house will give your broker their names as “strong buys.” He will leave a voice message that he has a hot recommendation that necessitates immediate action. You will be e-mailed a long list of names. You will buy one or two of them. Meanwhile, the manager in charge of your 401(k) retirement plan will be acquiring the entire list.

We can apply the reasoning to the selection of investment categories—as if they were the managers in the example above. Assume you are standing in 1900 with hundreds of investments to look at. There are the stock markets of Argentina, Imperial Russia, the United Kingdom, Unified Germany, and plenty of others to consider. A rational person would have bought not just the emerging country of the United States, but those of Russia and Argentina as well. The rest of the story is well-known; while many of the stock markets like those of the United Kingdom and the United States fared extremely well, the investor in Imperial Russia would have no better than medium-quality wallpaper in his hands. The countries that fared well are not a large segment of the initial cohort; randomness would be expected to allow a few investment classes to fare extremely well. I wonder if those “experts” who make foolish (and self-serving) statements like “markets will always go up in any twenty-year period” are aware of this problem.

COMPARATIVE LUCK

A far more acute problem relates to the outperformance, or the comparison, between two or more persons or entities. While we are certainly fooled by randomness when it comes to a single times series, the foolishness is compounded when it comes to the comparison between, say, two people, or a person and a benchmark. Why? Because both are random. Let us do the following simple thought experiment. Take two individuals, say, a person and his brother-in-law, launched through life. Assume equal odds for each of good and bad luck. Outcomes: lucky-lucky (no difference between them), unlucky-unlucky (again, no difference), lucky-unlucky (a large difference between them), unlucky-lucky (again, a large difference).

I recently attended for the first time a conference of investment managers where I sat listening to a very dull presenter comparing traders. His profession is to select fund managers and package them together for investors, something called “funds of funds” and I was listening to him as he was pouring out numbers on the screen. The first revelation was that I suddenly recognized the speaker, a former colleague biologically transformed by the passage of time. He used to be crisp, energetic, and nice; he became dull, portly, and inordinately comfortable with success. (He was not rich when I knew him—can people react to money in different ways? Do some take themselves seriously while others do not?) The second revelation was that while I suspected that he was fooled by randomness, the extent had to be far greater than one could imagine, particularly with the survivorship bias. A back of the envelope calculation showed that at least 97% of what he was discussing was just noise. The fact that he was comparing performances made the matter far worse.

Cancer Cures

When I return home from an Asian or European trip, my jet lag often causes me to rise at a very early hour. Occasionally, though very rarely, I switch on the TV set searching for market information. What strikes me in these morning explorations is the abundance of claims by the alternative medicine vendors of the curing power of their products. These no doubt are caused by the lower advertising rates at that time. To prove their claim, they present the convincing testimonial of someone who was cured thanks to their methods. For instance, I once saw a former throat cancer patient explaining how he was saved by a combination of vitamins for sale for the exceptionally low price of $14.95—in all likelihood he was sincere (although of course compensated for his account, perhaps with a lifetime supply of such medicine). In spite of our advances, people still believe in the existence of links between disease and cure based on such information, and there is no scientific evidence that can convince them more potently than a sincere and emotional testimonial. Such testimonial does not always come from the regular guy; statements by Nobel Prize winners (in the wrong discipline) could easily suffice. Linus Pauling, a Nobel Prize winner in chemistry, was said to believe in vitamin C’s medicinal properties, himself ingesting massive daily doses. With his bully pulpit, he contributed to the common belief in vitamin C’s curative properties. Many medical studies, unable to replicate Pauling’s claims, fell on deaf ears as it was difficult to undo the testimonial by a “Nobel Prize winner,” even if he was not qualified to discuss matters related to medicine.

Many of these claims have been harmless outside of the financial profits for these charlatans—but many cancer patients may have replaced the more scientifically investigated therapies, in favor of these methods, and died as a result of their neglecting more orthodox cures (again, the nonscientific methods are gathered under what is called “alternative medicine,” that is, unproven therapies, and the medical community has difficulties convincing the press that there is only one medicine and that alternative medicine is not medicine). The reader might wonder about my claims that the user of these products could be sincere, without it meaning that he was cured by the illusory treatment. The reason is something called “spontaneous remission,” in which a very small minority of cancer patients, for reasons that remain entirely speculative, wipe out cancer cells and recover “miraculously.” Some switch causes the patient’s immune system to eradicate all cancer cells from the body. These people would have been equally cured by drinking a glass of Vermont spring water or chewing on dried beef as they were by taking these beautifully wrapped pills. Finally, these spontaneous remissions might not be so

spontaneous; they might, at the bottom, have a cause that we are not yet sophisticated enough to detect.

The late astronomer Carl Sagan, a devoted promoter of scientific thinking and an obsessive enemy of nonscience, examined the cures from cancer that resulted from a visit to Lourdes in France, where people were healed by simple contact with the holy waters, and found out the interesting fact that, of the total cancer patients who visited the place, the cure rate was, if anything, lower than the statistical one for spontaneous remissions. It was lower than the average for those who did not go to Lourdes! Should a statistician infer here that cancer patients' odds of surviving deteriorates after a visit to Lourdes?

Professor Pearson Goes to Monte Carlo (Literally):

Randomness Does Not Look Random!

At the beginning of the twentieth century, as we were starting to develop techniques to deal with the notion of random outcomes, several methods were designed to detect anomalies. Professor Karl Pearson (father of Egon Pearson of Neyman-Pearson fame, familiar to every person who sat in a statistics 101 class) devised the first test of nonrandomness (it was in reality a test of deviation from normality, which, for all intents and purposes, was the same thing). He examined millions of runs of what was called a Monte Carlo (the old name for a roulette wheel) during the month of July 1902. He discovered that, with a high degree of statistical significance (with an error of less than one to a billion), the runs were not purely random. What! The roulette wheel was not random! Professor Pearson was greatly surprised at the discovery. But this result in itself tells us nothing; we know that there is no such thing as a pure random draw, for the outcome of the draw depends on the quality of the equipment. With enough minutiae one would be able to uncover the nonrandomness somewhere (e.g., the wheel itself may not have been perfectly balanced or perhaps the spinning ball was not completely spherical). Philosophers of statistics call this the reference case problem to explain that there is no true attainable randomness in practice, only in theory. Besides, a manager would question whether such nonrandomness can lead to any meaningful, profitable rules. If I need to gamble $1 on 10,000 runs and expect to make $1 for my efforts, then I would do much better in the part-time employment of a janitorial agency.

But the result bears another suspicious element. Of more practical relevance here is the following severe problem about nonrandomness. Even the fathers of statistical science forgot that a random series of runs need not exhibit a pattern to look random; as a matter of fact, data that is perfectly patternless would be extremely suspicious and appear to be man-made. A single random run is bound to exhibit some pattern—if one looks hard enough. Note that Professor Pearson was among the first scholars who were interested in creating artificial random data generators, tables one could use as inputs for various scientific and engineering simulations (the precursors of our Monte Carlo simulator). The problem is that they did not want these tables to exhibit any form of regularity. Yet real randomness does not look random!

I would further illustrate the point with the study of a phenomenon well-known as cancer clusters. Consider a square with 16 random darts hitting it with equal probability of being at any place in the square. If we divide the square into 16 smaller squares, it is expected that each smaller square will contain one dart on average—but only on average. There is a very small probability of having exactly 16 darts in 16 different squares. The average grid will have more than one dart in a few squares, and no dart at all in many squares. It will be an exceptionally rare incident that no (cancer) cluster would show on the grid. Now, transpose our grid with the darts in it to overlay a map of any region. Some newspaper will declare that one of the areas (the one with more than the average of darts) harbors radiation that causes cancer, prompting lawyers to start soliciting the patients.

The Dog That Did Not Bark: On Biases in Scientific Knowledge

By the same argument, science is marred by a pernicious survivorship bias, affecting the way research gets published. In a way that is similar to journalism, research that yields no result does not make it to print. That may seem sensible, as newspapers do not have to have a screaming headline saying that nothing new is taking place (though the Bible was smart enough to declare ein chadash tachat hashemesh—"nothing new under the sun," providing the information that things just do recur). The problem is that a finding of absence and an absence of findings get mixed together. There may be great information in the fact that nothing took place. As Sherlock Holmes noted in the Silver Blaze case—the curious thing was that the dog did not bark. More problematic, there are plenty of scientific results that are left out of publications because they are not statistically significant, but nevertheless provide information.

I HAVE NO CONCLUSION

I am frequently asked the question "When is it truly not luck?" There are professions in randomness for which performance is low in luck, like casinos, which manage to tame randomness. In finance? Perhaps. All traders are not speculative traders: There exists a segment called market makers whose job is to derive, like bookmakers, or even like store owners, an income against a transaction. If they speculate, their dependence on the risks of such speculation remains too small compared to their overall volume. They buy at a price and sell to the public at a more favorable one, performing large numbers of transactions. Such income provides them some insulation from randomness. Such category includes floor traders on the exchanges, bank traders who "trade against order flow," money changers in the souks of the Levant. The skills involved are sometimes rare to find: Fast thinking, alertness, a high level of energy, an ability to guess from the voice of the seller her level of nervousness; those who have them make a long career (that is, perhaps a decade).They never make it big, as their income is constrained by the number of customers, but they do well probabilistically. They are, in a way, the dentists of the profession.

Outside of this very specialized bookmaker-style profession, to be honest, I am unable to answer the question of who's lucky or unlucky. I can tell that person A seems less lucky than person B, but the confidence in such knowledge can be so weak as to be meaningless. I prefer to remain a skeptic. People frequently misinterpret my opinion. I never said that every rich man is an idiot and every unsuccessful person unlucky, only that in absence of much additional information it is preferable to reserve one's judgment. It is safer.