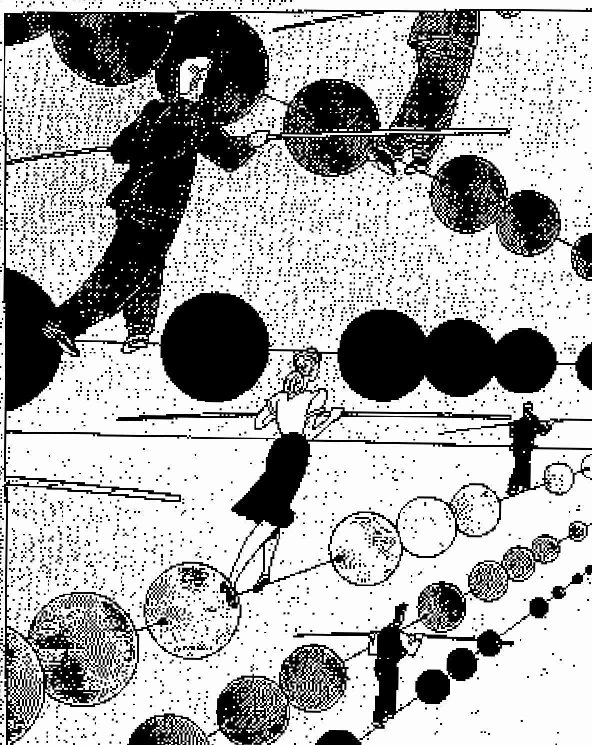
INNUMERACY

MATHEMATICAL ILLITERACY AND ITS CONSEQUENCES



'Our society would be unimaginably different if the average person truly understood the ideas in this marvellous and important little book' — **Douglas Hofstadter**

JOHN ALLEN PAULOS

2

Probability and Coincidence

It is no great wonder if, in the long process of time, while fortune takes her course hither and thither, numerous coincidences should spontaneously occur. —Plutarch

"You're a Capricorn, too. That's so exciting."

A man who travels a lot was concerned about the possibility of a bomb on board his plane. He determined the probability of this, found it to be low but not low enough for him, so now he always travels with a bomb in his suitcase. He reasons that the probability of two bombs being on board would be infinitesimal.

Some Birthday vs. a Particular Birthday

Sigmund Freud once remarked that there was no such thing as a coincidence. Carl Jung talked about the mysteries of synchronicity. People in general prattle ceaselessly about ironies here and ironies there. Whether we call them coincidences, synchronicities, or ironies, however, these occurrences are much more common than most people realize.

Some representative examples: "Oh, my brother-in-law went to school there, too, and my friend's son cuts the principal's birthdays are independent, the probability of two people not having March 19 as a birthday is 364/365 × 364/365. Thus, the probability of N people not having March 19 as a birthday is $(364/365)^N$, which, when N = 253, is approximately ½. Hence, the complementary probability that at least one of these 253 people was born on March 19 is also ½, or 50 percent.

The moral, again, is that some unlikely event is likely to occur, whereas it's much less likely that a particular one will. Martin Gardner, the mathematics writer, illustrates the distinction between general and specific occurrences by means of a spinner with the twenty-six letters of the alphabet on it. If the spinner is spun one hundred times and the letters recorded, the probability that the word CAT or WARM will appear is very small, but the probability of *some* word's appearing is high. Since I brought up the topic of astrology, Gardner's examples of the first letters of the names of the months and the planets are particularly appropriate. The months—JFMAMJJASOND—give us JASON; the planets—MVEMJSUNP—spell SUN. Significant? No.

The paradoxical conclusion is that it would be very unlikely for unlikely events not to occur. If you don't specify a predicted event precisely, there are an indeterminate number of ways for an event of that general kind to take place.

Medical quackery and television evangelism will be discussed in the next chapter, but it should be mentioned here that their predictions are usually sufficiently vague so that the probability of some event of the predicted kind occurring is very high; it's the particular predictions that seldom come true. That some nationally famous politician will undergo a sex-change operation, as a newspaper astrologer-psychic recently predicted, is considerably more likely than that New York's Mayor Koch will. That some viewer will be relieved of his gastric pains just as a television evangelist calls out the symptoms is considerably more likely than that a particular viewer will be. Likewise, insurance policies with broad coverage which compensates for any mishap are apt to be cheaper in the long run than insurance for a particular disease or a particular trip.

Chance Encounters

Two strangers from opposite sides of the United States sit next to each other on a business trip to Milwaukee and discover that the wife of one of them was in the tennis camp run by an acquaintance of the other's. This sort of coincidence is surprisingly common. If we assume each of the approximately 200 million adults in the United States knows about 1,500 people, and that these 1,500 people are reasonably spread out around the country, then the probability is about one in a hundred that they will have an acquaintance in common, and more than ninety-nine in a hundred that they will be linked by a chain of two intermediates.

We can be almost certain, then, given these assumptions, that two people chosen at random will be linked, as were the strangers on the business trip, by a chain of at most two intermediates. Whether they'll run down the 1,500 or so people they each know (as well as the acquaintances of each of these 1,500) during their conversation and thus become aware of the two intermediates linking them is another, more dubious matter.

These assumptions can be relaxed somewhat. Maybe the average adult knows fewer than 1,500 other adults, or, more likely, most of the people he or she does know live close by and are not spread about the country. Even in these cases, however, the probability of two randomly selected people being linked by two intermediates is unexpectedly high.

A more empirical approach to coincidental meetings was taken by psychologist Stanley Milgrim, who gave each member of a randomly selected group of people a document and a (different) "target individual" to whom the document was to be transmitted. The directions were that each person was to send the document to the person he knew who was most likely to know the target individual, and that he was to direct that person to do the same, until the target individual was reached. Milgrim found that the number of intermediate links ranged from two to ten, with five being the most common number. This study is

A Stock-Market Scam

Stock-market advisers are everywhere, and you can probably find one to say almost anything you might want to hear. They're usually assertive, sound quite authoritative, and speak a strange language of puts, calls, Ginnie Maes, and zero-coupons. In my humble experience, most don't really know what they're talking about, but presumably some do.

If from some stock-market adviser you received in the mail for six weeks in a row correct predictions on a certain stock index and were asked to pay for the seventh such prediction, would you? Assume you really are interested in making an investment of some sort, and assume further that the question is being posed to you before the stock crash of October 19, 1987. If you would be willing to pay for the seventh prediction (or even if you wouldn't), consider the following con game.

Some would-be adviser puts a logo on some fancy stationery and sends out 32,000 letters to potential investors in a stock index. The letters tell of his company's elaborate computer model, his financial expertise and inside contacts. In 16,000 of these letters he predicts the index will rise, and in the other 16,000 he predicts a decline. No matter whether the index rises or falls, a follow-up letter is sent, but only to the 16,000 people who initially received a correct "prediction." To 8,000 of them, a rise is predicted for the next week; to the other 8,000, a decline. Whatever happens now, 8,000 people will have received two correct predictions. Again, to these 8,000 people only, letters are sent concerning the index's performance the following week: 4,000 predicting a rise; 4,000, a decline. Whatever the outcome, 4,000 people have now received three straight correct predictions.

This is iterated a few more times, until 500 people have received six straight correct "predictions." These 500 people are now reminded of this and told that in order to continue to receive this valuable information for the seventh week they must each contribute $500. If they all pay, that's $250,000 for our adviser.

If this is done knowingly and with intent to defraud, this is an illegal con game. Yet it's considered acceptable if it's done unknowingly by earnest but ignorant publishers of stock newsletters, or by practitioners of quack medicine, or by television evangelists. There's always enough random success to justify almost anything to someone who wants to believe.

There is another quite different problem exemplified by these stock-market forecasts and fanciful explanations of success. Since they're quite varied in format and often incomparable and very numerous, people can't act on all of them. The people who try their luck and don't fare well will generally be quiet about their experiences. But there'll always be some people who will do extremely well, and they will loudly swear to the efficacy of whatever system they've used. Other people will soon follow suit, and a fad will be born and thrive for a while despite its baselessness.

There is a strong general tendency to filter out the bad and the failed and to focus on the good and the successful. Casinos encourage this tendency by making sure that every quarter that's won in a slot machine causes lights to blink and makes its own little tinkle in the metal tray. Seeing all the lights and hearing all the tinkles, it's not hard to get the impression that everyone's winning. Losses or failures are silent. The same applies to well-publicized stock-market killings vs. relatively invisible stock-market ruinations, and to the faith healer who takes credit for any accidental improvement but will deny responsibility if, for example, he ministers to a blind man who then becomes lame.

This filtering phenomenon is very widespread and manifests itself in many ways. Along almost any dimension one cares to choose, the average value of a large collection of measurements is about the same as the average value of a small collection, whereas the extreme value of a large collection is considerably more extreme than that of a small collection. For example, the average water level of a given river over a twenty-five-year period will be approximately the same as the average water level over a one-year period, but the worst flood over a twenty-five-year period is apt to be considerably higher than that over a one-year period. The average scientist in tiny Belgium will be comparable to the average scientist in the United States, even though the best scientist in the United States will in general be better than Belgium's best (we ignore obvious complicating factors and definitional problems).

So what? Because people usually focus upon winners and extremes whether they be in sports, the arts, or the sciences, there's always a tendency to denigrate today's sports figures, artists, and scientists by comparing them with extraordinary cases. A related consequence is that international news is usually worse than national news, which in turn is usually worse than state news, which is worse than local news, which is worse than the news in your particular neighborhood. Local survivors of tragedy are invariably quoted on TV as saying something like, "I can't understand it. Nothing like that has ever happened around here before."

One final manifestation: Before the advent of radio, TV, and film, musicians, athletes, etc., could develop loyal local audiences since they were the best that most of these people would ever see. Now audiences, even in rural areas, are no longer as satisfied with local entertainers and demand world-class talent. In this sense, these media have been good for audiences and bad for performers.

Expected Values: From Blood Testing to Chuck-a-Luck

Coincidences or extreme values catch the eye, but average or "expected" values are generally more informative. The expected value of a quantity is simply the average of its values weighted according to their probabilities. For example, if 1/4 of the time a quantity equals 2, 1/3 of the time it equals 6, another 1/3 of the time it equals 15, and the remaining 1/12 of the time it equals 54, then its expected value equals 12. This is so since $[12 = (2 \times 1/4) + (6 \times 1/3) + (15 \times 1/3) + (54 \times 1/12)]$.

As a simple illustration, consider a home-insurance company. Assume it has good reason to believe that, on average, each year one out of every 10,000 of its policies will result in a claim of $200,000; one out of 1,000 policies will result in a claim of $50,000; one out of 50 will result in a claim of $2,000; and the remainder will result in a claim of $0. The insurance company would like to know what its average payout is per policy written. The answer is the expected value, which in this case is ($200,000 × 1/10,000) + ($50,000 × 1/1,000) + ($2,000 × 1/50) + ($0 × 9,789/10,000) = $20 + $50 + $40 = $110.

The expected payout on a slot machine is determined in like manner. Each payout is multiplied by the probability of its occurring, and these products are then summed up to give the average or expected payout. For example, if cherries on all three dials result in a payout of $80 and the probability of this is $(1/20)^3$ (assume there are twenty entries on each dial, only one of which is a cherry), we multiply $80 by $(1/20)^3$ and then add to this product the products of the other payouts (a loss being considered a negative payout) and their probabilities.

An illustration which isn't quite so vanilla: Assume a medical clinic tests blood for a certain disease from which approximately one person in a hundred suffers. People come to the clinic in groups of fifty, and the director wonders whether, instead of testing them individually, he should pool the fifty samples and test them all together. If the pooled sample is negative, he could pronounce the whole group healthy, and if not, he could then test each person individually. What is the expected number of tests the director will have to perform if he pools the blood samples?

The director will have to perform either one test (if the pooled sample is negative) or fifty-one tests (if it's positive). The probability that any one person is healthy is 99/100, and so the probability that all fifty people are healthy is $(99/100)^{50}$. Thus, the probability that he'll have to perform just one test is $(99/100)^{50}$. On the other hand, the probability that at least one person suffers from the disease is the complementary probability $[1 - (99/100)^{50}]$, and so the probability of having to perform fifty-one tests is

$[1 - (99/100)^{50}]$. Thus, the expected number of tests necessary is $(1 \text{ test} \times (99/100)^{50}) + (51 \text{ tests} \times [1 - (99/100)^{50}])$ = approximately 21 tests.

If there are large numbers of people having the blood test, the medical director would be wise to take part of each sample, pool it, and test this pooled sample first. If necessary, he could then test the remainders of each of the fifty samples individually. On average, this would require only twenty-one tests to test fifty people.

An understanding of expected values is helpful in analyzing most casino games, as well as the lesser-known game of chuck-a-luck which is played at carnivals in the Midwest and England.

The spiel that goes with chuck-a-luck can be very persuasive. You pick a number from 1 to 6 and the operator rolls three dice. If the number you pick comes up on all three dice, the operator pays you $3; if it comes up on two of the three dice, he pays you $2; and if it comes up on just one of the three dice, he pays you $1. Only if the number you picked doesn't come up at all do you pay him anything—just $1. With three different dice, you have three chances to win, and furthermore you'll sometimes win more than $1, while that is your maximum loss.

As Joan Rivers might say, "Can we calculate?" (If you'd rather not calculate, skip to the end of this section.) The probability of your winning is clearly the same no matter what number you choose, so, to make the calculation specific, assume you always pick the number 4. Since the dice are independent, your chances that a 4 will come up on all three dice are $1/6 \times 1/6 \times 1/6 = 1/216$; so, approximately 1/216th of the time you'll win $3.

Your chances of a 4 coming up only twice are a little harder to calculate unless you use the binomial probability distribution mentioned in Chapter 1, which I'll derive again in this context. A 4 coming up on two of the three dice can happen in three different and mutually exclusive ways: X44, 4X4, or 44X, the X indicating a non-4. The probability of the first is $5/6 \times 1/6 \times 1/6 = 5/216$, a result which holds true for the second and third ways as well. Adding, we find that the probability of a 4 coming up on two of the three dice is $15/216$, which is the fraction of the time you'll win $2.

The probability of obtaining exactly one 4 among the three dice is likewise determined by breaking the event into the three mutually exclusive ways it can happen. The probability of obtaining 4XX is $1/6 \times 5/6 \times 5/6 = 25/216$, which is also the probability of obtaining X4X and XX4. Adding, we get $75/216$ for the probability of exactly one 4 coming up on the three dice, and hence for the probability of your winning $1. To find the probability that no 4s come up when we roll three dice, we find how much probability is left over. That is, we subtract $(1/216 + 15/216 + 75/216)$ from 1 (or 100%) to get $125/216$. Thus, on the average, 125 out of 216 times you play chuck-a-luck, you'll lose $1.

The expected value of your winnings is thus $(\$3 \times 1/216) + (\$2 \times 15/216) + (\$1 \times 75/216) + (-\$1 \times 125/216) = \$(-17/216) = -\$.08$; and so, on the average, you would lose approximately eight cents every time you played this seemingly attractive game.

Choosing a Spouse

There are two approaches to love—through the heart and through the head. Neither one seems to work very well alone, but together . . . they still don't work too well. Nevertheless, there's probably a better chance of success if both are used. Upon thinking of past loves, someone who approaches romance through the heart is likely to bemoan lost opportunities and conclude that he or she will never again love as deeply. Someone who takes a more hardheaded approach may be interested in the following result in probability.

The model we'll consider assumes that our heroine—call her Myrtle—has reason to believe that she'll meet N potential spouses (spice?) during her "dating life." N could be two for some women, two hundred for others. The question Myrtle poses to herself is: When should I accept Mr. X and forgo the suitors who would come after him, some of whom may possibly be "better" than he? We'll assume she meets men sequentially, can judge the relative suitability for her of those she's met, and once she's rejected someone, he's gone forever.

For illustration, suppose Myrtle has met six men so far and that she rates them as follows: 3 5 1 6 2 4. That is, of the six men she's met, she liked the first one she met third-best, the second one she liked fifth-best, the third one she liked best of all, and so on. If the seventh man she meets she prefers to everyone except her favorite, her updated ranking would become: 4 6 1 7 3 5 2. After each man, she updates her relative ranking of her suitors and wonders what rule she should follow in order to maximize her chances of choosing the best of her projected N suitors.

The derivation of the best policy uses the idea of conditional probability (which we'll introduce in the next chapter) and a little calculus. The policy itself, though, is quite simple to describe. Call a suitor a heartthrob if he's better than all previous candidates. Myrtle should reject approximately the first 37 percent of the N candidates she's likely to meet, and then accept the first suitor after that (if any) who is a heartthrob.

For instance, suppose Myrtle isn't overly attractive and is likely to meet only four eligible suitors, and suppose further that these four men are equally likely to come to her in any of the twenty-four possible orderings ($24 = 4 \times 3 \times 2 \times 1$).

Since 37 percent is between 25 percent and 50 percent, the policy is ambiguous here, but the two best strategies correspond to the following: (A) Pass up the first candidate (25 percent of $N = 4$) and accept the first heartthrob after that. (B) Pass up the first two candidates (50 percent of $N = 4$) and accept the first heartthrob after that. Strategy A will result in Myrtle's choosing the best suitor in eleven of the twenty-four instances, while strategy B will result in success in ten of the twenty-four instances.

The list of all such sequential orderings is below, with the number 1 representing, as before, the suitor Myrtle would most prefer, 2 her second choice, etc. Thus, the ordering 3 2 1 4 indicates that she meets her third choice first, her second choice second, her first choice she meets third, and her last choice last. The orderings are marked with an A or a B to indicate in which instances these strategies result in her getting her first choice.

1234 · 1243 · 1324 · 1342 · 1423 · 1432 · 2134(A) · 2143(A) · 2314(A,B) · 2341(A,B) · 2413(A,B) · 2431(A,B) · 3124(A) · 3142(A) · 3214(B) · 3241(B) · 3412(A,B) · 3421 · 4123(A) · 4132(A) · 4213(B) · 4231(B) · 4312(B) · 4321

If Myrtle is quite attractive and can expect to have twenty-five suitors, her best strategy would still be to reject the first nine of these suitors (37 percent of 25) and then accept the first heartthrob after that. This could be verified by tabulation, as above, but the tables get unwieldy and it's best to accept the general proof. (Needless to say, the same analysis holds if the person seeking a spouse is a Mortimer and not a Myrtle.)

For large values of N, the probability that Myrtle will find her Mr. Right following this 37 percent rule is also approximately 37 percent. Then comes the hard part: living with Mr. Right. Variants of the model exist with more romantically plausible constraints.

Coincidence and the Law

In 1964 in Los Angeles a blond woman with a ponytail snatched a purse from another woman. The thief fled on foot but was later spotted entering a yellow car driven by a black man with a beard and a mustache. Police investigation eventually discovered a blond woman with a ponytail who regularly associated with a bearded and mustachioed black man who owned a yellow car. There wasn't any hard evidence linking the couple to the crime, or any witnesses able to identify either party. There was, however, agreement on the above facts.

The prosecutor argued that the probability was so low that such a couple existed that the police investigation must have turned up the actual culprits. He assigned the following probabilities to the characteristics in question: yellow car—1/10; man with a mustache—1/4; woman with a ponytail—1/10; woman with blond hair—1/3; black man with a beard—1/10; interracial couple in a car—1/1,000. The prosecutor further argued that the characteristics were independent, so that the probability that a randomly selected couple would have all of them would be 1/10 × 1/4 × 1/10 × 1/3 × 1/10 × 1/1,000 = 1/12,000,000, a number so low the couple must be guilty. The jury convicted them.

The case was appealed to the California supreme court, where it was overturned on the basis of another probability argument. The defense attorney in that trial argued that 1/12,000,000 was not the relevant probability. In a city the size of Los Angeles, with maybe 2,000,000 couples, the probability was not that small, he maintained, that there existed more than one couple with that particular list of characteristics, given that there was at least one such couple—the convicted couple. On the basis of the binomial probability distribution and the 1/12,000,000 figure, this probability can be determined to be about 8 percent—small, but certainly allowing for reasonable doubt. The California supreme court agreed and overturned the earlier guilty verdict.

Whatever the problems of the one in 12,000,000 figure, rarity by itself shouldn't necessarily be evidence of anything. When one is dealt a bridge hand of thirteen cards, the probability of being dealt that particular hand is less than one in 600 billion. Still, it would be absurd for someone to be dealt a hand, examine it carefully, calculate that the probability of getting it is less than one in 600 billion, and then conclude that he must not have been dealt that very hand because it is so very improbable.

In some contexts, improbabilities are to be expected. Every bridge hand is quite improbable. Likewise with poker hands or lottery tickets. In the case of the California couple, improbability carries more weight, but still, their defense attorney's argument was the right one.

Why is it, incidentally, if all the 3,838,380 ways of choosing six numbers out of forty are equally likely, that a lottery ticket with the numbers 2 13 17 20 29 36 is for most people much preferable to one with the numbers 1 2 3 4 5 6? This is, I think, a fairly deep question.

The following sports anomaly has legal implications as well. Consider two baseball players, say, Babe Ruth and Lou Gehrig. During the first half of the season, Babe Ruth hits for a higher batting average than Lou Gehrig. And during the second half of the season, Babe Ruth again hits for a higher average than Lou Gehrig. But for the season as whole, Lou Gehrig has a higher batting average than Babe Ruth. Could this be the case? Of course, the mere fact that I pose the question may cause some misgivings, but at first glance such a situation seems impossible.

What can happen is that Babe Ruth could hit .300 the first half of the season and Lou Gehrig only .290, but Ruth could bat two hundred times to Gehrig's one hundred times. During the second half of the season, Ruth could bat .400 and Gehrig only .390, but Ruth could come to bat only a hundred times to Gehrig's two hundred times at bat. The result would be a higher overall batting average for Gehrig than for Ruth: .357 vs. .333. You can't average batting averages.

There was an intriguing discrimination case in California several years ago which had the same formal structure as this batting-average puzzle. Looking at the proportion of women in graduate school at a large university, some women filed a lawsuit claiming that they were being discriminated against by the graduate school. When administrators sought to determine which departments were most guilty, they found that in each department a higher percentage of women applicants were admitted than men applicants. Women, however, applied in disproportionately large numbers to departments such as English and psychology that admitted only a small percentage of their applicants, whereas men applied in disproportionately large numbers to departments such as mathematics and engineering that admitted a much higher percentage of their applicants. The men's admissions pattern was analogous to Gehrig's hitting pattern—coming to bat more often during the second half of the season when getting a hit is easier.

Another counter-intuitive problem involving seemingly disproportionate probabilities concerns a New York City man who has a woman friend in the Bronx and one in Brooklyn. He is equally attached (or perhaps unattached) to each of them and thus is indifferent to whether he catches the northbound subway to the Bronx or the southbound subway to Brooklyn. Since both trains run every twenty minutes throughout the day, he figures he'll let the subway decide whom he'll visit, and take the first train which comes along. After a while, though, his Brooklyn woman friend, who's enamored of him, begins to complain that he shows up for only about one-fourth of his dates with her, while his Bronx friend, who's getting sick of him, begins to complain that he appears for three-fourths of his dates with her. Aside from callowness, what is this man's problem?

The simple answer follows, so don't read on if you want to think a bit. The man's more frequent trips to the Bronx are a result of the way the trains are scheduled. Even though they each come every twenty minutes, the schedule may be something like the following: Bronx train, 7:00; Brooklyn train, 7:05; Bronx train, 7:20; Brooklyn train, 7:25; and so on. The gap between the last Brooklyn train and the next Bronx train is fifteen minutes, three times as long as the five-minute gap between the last Bronx train and the next Brooklyn train, and thus accounts for his showing up for three-fourths of his dates in the Bronx and only for one-fourth of his Brooklyn dates.

Countless similar oddities result from our conventional ways of measuring, reporting, and comparing periodic quantities, whether they be the monthly cash flow of a government or the regular daily fluctuations in body temperature.

Fair Coins and Life's Winners and Losers

Imagine flipping a coin many times in succession and obtaining some sequence of heads and tails; say, HHTHTTHH THTTTHTTHHHTHTTHHTHHTTHTHHTTHHTHT HHHHTHHHTT. If the coin is fair, there are a number of extremely odd facts about such sequences. For example, if one were to keep track of the proportion of the time that the number of heads exceeded the number of tails, one might be surprised that it is rarely close to half.

Imagine two players, Peter and Paul, who flip a coin at the rate of once a day and who bet on heads and tails respectively. Peter is ahead at any given time if there've been more heads up until then, while Paul is ahead if there've been more tails. Peter and Paul are each equally likely to be ahead at any given time, but whoever is ahead will probably have been ahead almost the whole time. If there have been one thousand coin flips, then if Peter is ahead at the end, the chances are considerably greater that he's been ahead more than 90 percent of the time, say, than that he's been ahead between 45 percent and 55 percent of the time! Likewise, if Paul is ahead at the end, it's considerably more likely that he's been ahead more than 96 percent of the time than that he's been ahead between 48 percent and 52 percent of the time.

Perhaps the reason this result is so counter-intuitive is that most people tend to think of deviations from the mean as being somehow bound by a rubber band: the greater the deviation, the greater the restoring force toward the mean. The so-called gambler's fallacy is the mistaken belief that because a coin has come up heads several times in a row, it's more likely to come up tails on its next flip (similar notions hold for roulette wheels and dice).

The coin, however, doesn't know anything about any mean or rubber band, and if it's landed heads 519 times and tails 481 times, the difference between its heads total and its tails total is just as likely to grow as to shrink. This is true despite the fact that the proportion of heads does approach ½ as the number of coin flips increases. (The gambler's fallacy should be distinguished from another phenomenon, regression to the mean, which is valid. If the coin is flipped a thousand more times, it is more likely than not that the number of heads on the second thousand flips would be smaller than 519.)

In terms of ratios, coins behave nicely: the ratio of heads to tails gets closer to 1 as the number of flips grows. In terms of absolute numbers, coins behave badly: the difference between the number of heads and the number of tails tends to get bigger as we continue to flip the coin, and the changes in lead from head to tail or vice versa tend to become increasingly rare.

If even fair coins behave so badly in an absolute sense, it's not surprising that some people come to be known as "losers" and others as "winners" though there is no real difference between them other than luck. Unfortunately perhaps, people are more sensitive to absolute differences between people than they are to rough equalities between them. If Peter and Paul have won, respectively, 519 and 481 trials, Peter will likely be termed a winner and Paul a loser. Winners (and losers) are often, I would guess, just people who get stuck on the right (or wrong) side of even. In the case of coins, it can take a long, long time for the lead to switch, longer often than the average life.

The surprising number of consecutive runs of heads or tails of various lengths give rise to further counter-intuitive notions. If Peter and Paul flip a fair coin every day to determine who pays for lunch, then it's more likely than not that at some time within about nine weeks Peter will have won five lunches in a row, as will have Paul. And at some period within about five to six years it's likely that each will have won ten lunches in a row.

Most people don't realize that random events generally can seem quite ordered. The following is a computer printout of a random sequence of Xs and Os, each with probability ½.

O X X X O O O X X X O X X X O X X X X O O X X O X X
O X O O X O X O O O O X O X X O O O X X X O X O X X

X X X X X X X O X X X O X O X X X X O X O O X X X O
O O X X X X X O O X X O O O X X O O O O O X X O O X
X X X X X O X X X X O O X X X X O O X X O X X O O X
X O X O X O O X X X O X X O X X X X O X X O X X X X
X X X X X O X X X X X O O O O O X O O X X X O O X X
X X O O X O O X O X X X O X X X X O O O O X O X O X
X O X X X O O X X O O O O X X X X X O O O O X X X X
O X X O O X X X X X X O X X O O O O O O O X O X X X
X X O O O X X O X X X O O O O X O X O X O O X X X X
O X O X X X O X X O O X X O X O O X O O X X X O X X

Note the number of runs and the way there seem to be clumps and patterns. If we felt compelled to account for these, we would have to invent explanations that would of necessity be false. Studies have been done, in fact, in which experts in a given field have analyzed such random phenomena and come up with cogent "explanations" for the patterns.

With this in mind, think of some of the pronouncements of stock analysts. The daily ups and downs of a particular stock, or of the stock market in general, certainly aren't completely random as the Xs and Os above are, but it's safe to say that there is a very large element of chance involved. You might never guess this, however, from the neat post hoc analyses that follow each market's close. Commentators always have a familiar cast of characters to which they can point to explain any rally or any decline. There's always profit-taking or the federal deficit or something or other to account for a bearish turn, and improved corporate earnings or interest rates or whatever to account for a bullish one. Almost never does a commentator say that the market's activity for the day or even the week was largely a result of random fluctuations.

The Hot Hand and the Clutch Hitter

The clumps, runs, and patterns that random sequences evince can to an extent be predicted. Sequences of heads and tails of a given length, say twenty flips, generally have a certain number of consecutive runs of heads. A sequence of twenty coin flips which resulted in ten heads followed by ten tails (HHHHHHHHHHTTTTTTTTTT) is said to have just one run of heads. A sequence of twenty coin flips which resulted in heads and tails alternating (HTHTHTHTHTHTHTHTHTHT) is said to have ten runs of heads. Both these sequences are unlikely to be randomly generated. A sequence of twenty flips with six runs of heads (say, HHTHHTHTTHHHTTHHTTHT) is more likely to have been generated at random.

Criteria like this can be used to determine how likely it is that sequences of heads and tails or Xs and Os or hits and misses are randomly generated. In fact, psychologists Amos Tversky and Daniel Kahneman have analyzed the sequences of hits and misses of professional basketball players whose basket-making was about 50 percent and found that they seemed to be completely random—that a "hot hand" in basketball, one that would result in an inordinate number of long streaks (runs) of consecutive baskets, just didn't seem to exist. The streaks that did occur were most likely due to chance. If a player attempts twenty shots per night, for example, the probability is surprisingly almost 50 percent that he will hit at least four straight baskets sometime during the game. There's a 20 percent to 25 percent probability that he will achieve a streak of at least five straight baskets sometime during the game, and approximately a 10 percent chance that he will have a streak of six or more consecutive baskets.

Refinements can be made when the shooting percentage is other than 50 percent, and similar results seem to hold. A player who scores 65 percent of his shots, say, scores his points in the way a biased coin which lands heads 65 percent of the time "scores" its heads; i.e., each shot is independent of the last.

I've always suspected that notions like a "hot hand" or a "clutch hitter" or a "team that always comes back" were exaggerations used by sportswriters and sportscasters just to have something to talk about. There surely is something to these terms, but too often they're the result of minds intent on discovering meaning where there is only probability.

A very long hitting streak in baseball is a particularly amazing sort of record, so unlikely as to seem virtually unachievable and almost immune to probabilistic prediction. Several years ago Pete Rose set a National League record by hitting safely in forty-four consecutive games. If we assume for the sake of simplicity that he batted .300 (30 percent of the time he got a hit, 70 percent of the time he didn't) and that he came to bat four times per game, the chances of his not getting a hit in any given game were, assuming independence, $(.7)^4 = .24$. (Remember, independence means he got hits in the same way a coin which lands heads 30 percent of the time gets heads.) So the probability he would get at least one hit in any given game was $1 - .24 = .76$. Thus, the chances of his getting a hit in any given sequence of forty-four consecutive games were $(.76)^{44} = .0000057$, a tiny probability indeed.

The probability that he would hit in a consecutive string of exactly forty-four games at some time during the 162-game season is higher—.000041 (determined by adding up the ways in which he might hit safely in some string of exactly forty-four consecutive games, and ignoring the negligible probability of more than one streak). The probability that he'd hit safely in at least forty-four straight games is about four times higher still. If we multiply this latter figure by the number of players in the Major Leagues (adjusting the figure drastically downward for lower batting averages) and then multiply by the approximate number of years there has been baseball (adjusting for the various numbers of players in different years), we determine that it's actually not unlikely that some Major Leaguer should at some time have hit safely in at least forty-four consecutive games.

One last remark: I've examined Rose's streak of forty-four games rather than DiMaggio's seemingly more impressive streak of fifty-six games because, given the differences in their respective batting averages, Rose's streak was a slightly more unlikely accomplishment (even with Rose's longer season of 162 games).

Rare events such as batting streaks that are the result of chance are not individually predictable, yet the pattern of their occurrence is probabilistically describable. Consider a more prosaic kind of event. One thousand married couples who desire to have three children each are tracked for ten years, during which time 800 of them, assume, do manage to produce three children. The probability any given couple has three girls is $\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$, so approximately a hundred of these 800 couples will have three girls each. By symmetry, about a hundred of the couples will have three boys each. There are three different sequences in which a family might have two girls and a boy—GGB, GBG, or BGG, where the order of the letters indicates birth order—and each of the three sequences has the same probability of $\frac{1}{8}$, or $(\frac{1}{2})^3$. Thus, the probability of having two girls and a boy is $\frac{3}{8}$, and so approximately 300 of the 800 couples will have such a family. By symmetry, about 300 couples will also have two boys and a girl.

Nothing is very surprising about the above, but the same sort of probabilistic description (utilizing mathematics slightly more difficult than the above binomial distribution) is possible with very rare events. The number of accidents each year at a certain intersection, the number of rainstorms per year in a given desert, the number of cases of leukemia in a specified county, the annual number of deaths due to horse kicks among certain cavalry units of the Prussian Army have all been described quite accurately by the so-called Poisson probability distribution. It's necessary first to know roughly how rare the event is. But if you do know, you can use this information along with the Poisson formula to get a quite accurate idea of, for example, in what percentage of years there would be no deaths due to horse kicks, in what percentage of years there would be one such death, in what percentage of years two, in what percentage three, and so on. Likewise you could predict the percentage of years in which there would be no desert rainstorms, one such storm, two storms, three, and so on.

In this sense, even very rare events are quite predictable.