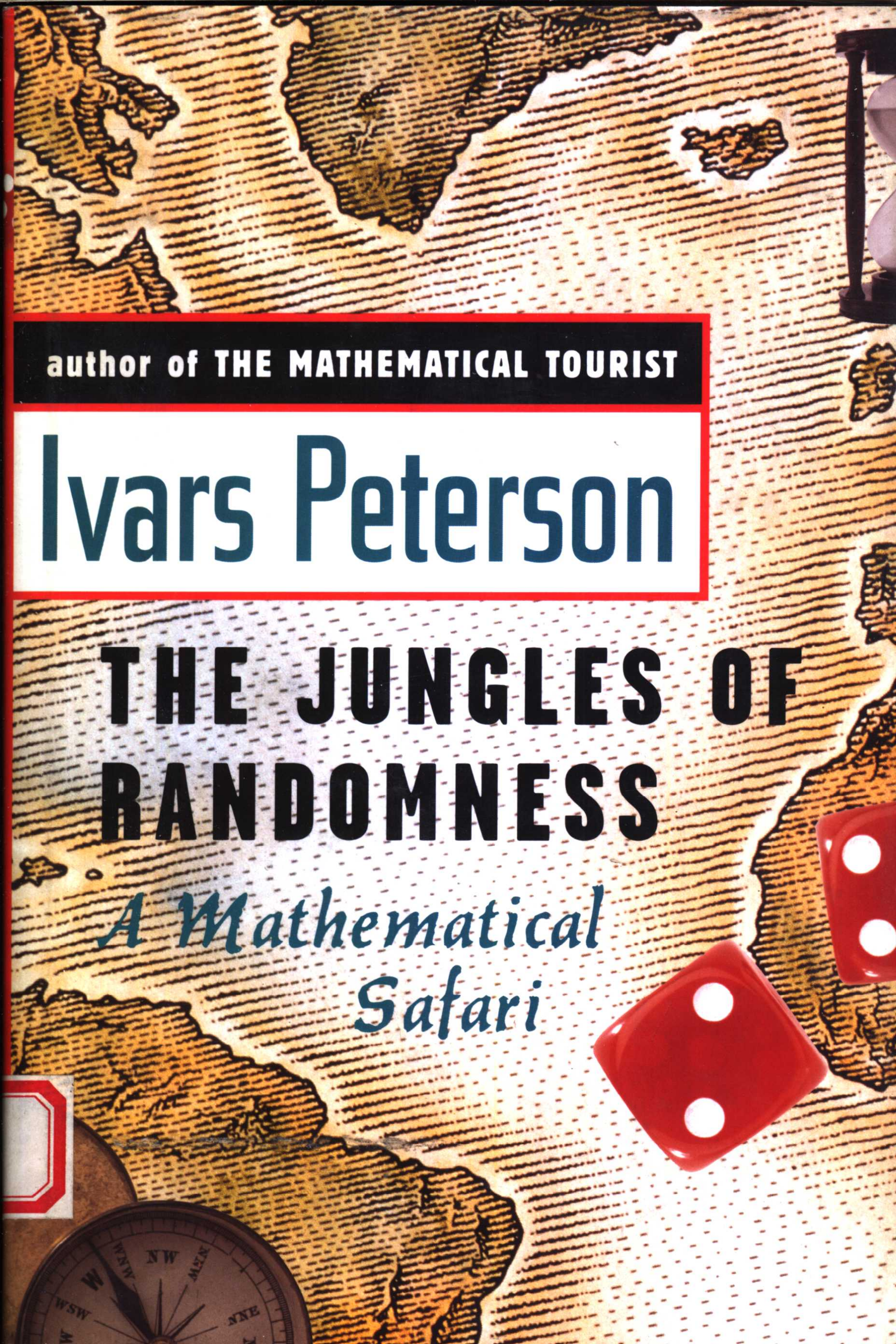

author of THE MATHEMATICAL TOURIST

Ivars Peterson

THE JUNGLES OF RANDOMNESS

A Mathematical Safari

The Jungles of Randomness

A Mathematical Safari

Ivars Peterson

John Wiley & Sons, Inc.
New York ▷ Chichester ▷ Weinheim ▷ Brisbane ▷ Singapore ▷ Toronto

This text is printed on acid-free paper.

Published by John Wiley & Sons, Inc.

Library of Congress Cataloging-in-Publication Data
Peterson, Ivars.
The jungles of randomness: a mathematical safari
Ivars Peterson.
p. cm.
Includes bibliographical references and index.
ISBN 0-471-16449-6 (alk. paper)
1. Probabilities—Popular works. I. Title.
QA273.15.P48 1997
519.2—dc21 97-1275
Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

Preface
Infinite Possibility

All mimsy were the borogoves,
And the mome raths outgrabe.

—*Lewis Carroll (1832–1898),* "Jabberwocky"

An ape sits hunched over a keyboard. A long hairy finger bangs a key, and the letter *a* appears on the computer screen. Another random stab produces *n*, then a space, then *a*, *p*, and *e*. That an ape would generate this particular sequence of characters is, of course, highly improbable. In the realm of random processes, however, any conceivable sequence of characters is possible, from utter gibberish to the full text of this book.

The seemingly infinite possibilities offered by randomness have long intrigued me. Years ago when I was a high school student, I came across a provocative statement by Arthur Stanley Eddington (1882–1944), a prominent astronomer and physicist. "If an army of monkeys were strumming on typewriters, they *might* write all the books in the British Museum," he noted.

Eddington wanted to emphasize the improbability of such an outcome, and his remark was meant as an example of something that could happen in principle but never in practice. I was left with the indelible image of a horde of monkeys forever pecking away at typewriters, generating the world's literature.

The preface that you are now reading contains nearly two thousand words, or roughly ten thousand characters, including spaces. An ape at a keyboard can choose among twenty-six letters and about fourteen other keys for punctuation, numbers, and spaces. Thus, it has one chance in forty of hitting *a* as the first letter, one chance in forty of picking *n* next, and so on. To write the entire preface, the ape would have to make the correct choice again and again. The probability of such an occurrence is one in forty multiplied by itself ten thousand times, or one in $40^{10,000}$. That figure is overwhelmingly larger than the estimated number of atoms in the universe, which is a mere 10^{80}.

One would have to wait an exceedingly long time before a member of a troop of apes happened to compose this book by chance, let alone the millions of volumes in the Library of Congress and the British Museum. Sifting through the troop's vast output to find the flawless gems, including original works of significant merit, would itself be a notably frustrating, unrewarding task. By eschewing randomness, a human author, on the other hand, can generate a meaningful string of characters far more efficiently than an ape, and the output generally requires considerably less editing.

Most people, including mathematicians and scientists, would say that they have a good idea of what the word *random* means. They can readily give all sorts of examples of random processes, from the flipping of a coin to the decay of a radioactive atomic nucleus. They can also list phenomena in which chance doesn't appear to play a role, from the motion of Earth around the sun to the ricochets of a ball between the cushions of a billiard table and the steady vibrations of a violin's plucked string.

Often, we use the word *random* loosely to describe something that is disordered, irregular, patternless, or unpredictable. We link it with chance, probability, luck, and coincidence. However, when we examine what we mean by *random* in various contexts, ambiguities and uncertainties inevitably arise. Tackling the subtleties of randomness allows us to go to the root of what we can understand of the universe we inhabit and helps us to define the limits of what we can know with certainty.

We think of flipping a coin as a way of making a blind choice, yet in the hands of a skilled magician the outcome may be perfectly predictable. Moreover, a process governed entirely by chance can lead to a completely ordered result, whether in the domain of monkeys pounding on keyboards or atoms locking into place to form a crystal. At the same time, a deterministic process can produce an unpredictable outcome, as seen in the waywardness of a ball rebounding within the confines of a stadium-shaped billiard table or heard in the screeches of an irregularly vibrating violin string. We can even invent mathematical formulas to generate predictable sequences of numbers that computers can then use to simulate the haphazard wanderings of perfume molecules drifting through the air.

It's useful to distinguish between a random process and the results of such a process. For example, we think of typing monkeys as generators of random strings of characters. If we know that such a random process is responsible for a given string, we may be justified in labeling or interpreting the string itself as random. However, if we don't know the source of a given string, we are forced to turn to other methods to determine what, if anything, the string means. Indeed, reading itself involves just such a search for meaning among the lines of characters printed on a page or displayed on a computer screen.

Consider the passage from Lewis Carroll's poem "Jabberwocky" that starts off the preface. From the presence of a few familiar words, the pattern of spaces, and the vowel-consonant structure of the remaining words, we would surmise that the author intended those lines to mean something, even though we don't understand many of the words. If the same passage were to come from typing monkeys, however, we might very well reject it as gibberish, despite the fragments of structure and pattern.

Similarly, in flipping a coin we know from experience (or theory) that we're likely to obtain an equal number of heads and tails in a long sequence of tosses. So if we see twenty-five heads in a row, it might be the legitimate though improbable result of a random process. However, it might also be advisable to check whether the coin is fair and to find out something about the fellow who's doing the flipping. The context determines how we interpret the data.

At the same time, just because we happen to see a sequence of roughly equal numbers of heads and tails doesn't mean that the results arise from tosses of a fair coin. It's possible to program a computer with a numerical recipe that involves no randomness yet gives the same distribution of heads and tails. Thus, given an arbitrary sequence of heads and tails, there's really no way to tell with confidence whether it's the result of a random process or it's been generated by a formula based on simple arithmetic.

"From a purely *operational* point of view . . . the concept of randomness is so elusive as to cease to be viable," the mathematician Mark Kac said in a 1983 essay on the nature of randomness. Kac also took a critical look at the different ways in which we sometimes interpret randomness in different contexts. For example, in the book *Chance and Necessity,* the biologist Jacques Monod (1910–1976) suggested that a distinction be made between "disciplined" chance, as used in physics to describe, say, radioactive decay, and "blind" chance. As an example of the latter, he cited the death of a doctor who, on his way to see a patient, was killed by a brick that fell from the roof of a building.

Kac argued that the distinction Monod makes really isn't meaningful. Although statistics on doctors killed by falling bricks aren't readily available, there are extensive data on Prussian soldiers kicked to death by horses—events that also fall under the category of blind chance. When one compares data on the number of soldiers killed in specified time intervals with data on the number of radioactive decays that have occurred in analogous periods, the two distributions of events look very similar.

Mathematics and statistics provide ways to sort through the various meanings of randomness and to distinguish between what we can and cannot know. They help us shape our expectations in different situations. In many cases, we find that there are no guarantees, only probabilities. We need to learn to recognize such limitations on certainty.

The Jungles of Randomness offers a random trek through the mélange of order and disorder that characterizes everyday experience. Along the way, my intention is to reveal a little of the immense, though often overlooked, impact of mathematics on our lives, to examine its power to explain, to suggest its austere elegance and beauty, and to provide a glimmer of its fundamental playfulness.

The search for pattern is a pervasive theme in mathematics. It is this pursuit that brings to our attention the curious interplay of order hidden in randomness and the randomness that is embedded in order. It's part of what makes mathematics such an alluring sport for mathematicians.

My aim is to provide a set of mathematical X rays that disclose the astonishing scope of randomness. The mathematical skeletons unveiled in these revealing snapshots serve as a framework for understanding a wide range of phenomena, from the vagaries of roulette wheels to the synchronization of cells in a beating heart. It's like opening up a watch to see what makes it tick. Instead of gears, levers, and wheels, however, we see equations and other pieces of mathematical apparatus.

Characterizing the vibrations of a drum's membrane, arranging points on the surface of a sphere, modeling the synchronized blink of a cloud of fireflies in Thailand, and playing games of chance are among the mathematical pastimes that provide connections to various aspects of everyday life. Each of those playful activities has prompted new thinking in mathematics. Each one brings randomness into play.

Mathematics encompasses the joy of solving puzzles, the exhilaration of subduing stubborn problems, the thrill of discerning patterns and making sense of apparent nonsense, and the immense satisfaction of nailing down an eternal truth. It is above all a human enterprise, one that is sometimes pursued simply for its own sake with nary a practical application in mind and sometimes inspired by a worldly concern but invariably pushed into untrodden territory. Mathematical research continually introduces new ideas and uncovers intriguing connections between old, well-established notions. Chance observations and informed guesses develop into entirely new fields of inquiry. Almost miraculously, links to the rest of the world inevitably follow.

With its system of theorems, proofs, and logical necessity, mathematics offers a kind of certainty. The tricky part lies in establishing meaningful connections between the abstract mathematical world that we create in our minds and the everyday world in which we live. When we find such links, mathematics can deliver accurate descriptions, yield workable solutions to real-world problems, and generate precise predictions. By making connections, we breathe life into the abstractions and symbols of the mathematicians' games.

Intriguingly, the mathematics of randomness, chaos, and order also furnishes what may be a vital escape from absolute certainty—an opportunity to exercise free will in a deterministic universe. Indeed, in the interplay of order and disorder that makes life interesting, we appear perpetually poised in a state of enticingly precarious perplexity. The universe is neither so crazy that we can't understand it at all nor so predictable that there's nothing left for us to discover.

So, the trek through the jungles of randomness starts with games of chance. It proceeds across the restless sea of life, from the ebb and flow of human concourse to the intricacies of biological structure and the dynamics of flashing fireflies. It wanders into the domain of sounds and oscillations and the realm of fractals and noise. It emerges from the territory of complete chaos as a random walk. Glimpses of gambling lead to a lifetime of chance.

Let the games begin!

Contents

1

The Die Is Cast

"Iacta alea est!"

—*Julius Caesar (100–44 B.C.)*

A die tumbles out of a cupped hand, bounces on the carpet a few times, rolls a short distance, then teeters to a stop. The uppermost face of the white cube shows four black dots arranged in a square.

Grinning, a child briskly moves a red token four squares to the right along the bottom row of a large checkerboard grid. The token lands on a square marked with the foot of a ladder. The player immediately takes the shortcut, advancing the token up the rungs to a higher row. Just ahead lies a square ominously marked with the upper end of a chute—the start of a costly detour.

With moves governed entirely by the roll of a die, Chutes and Ladders is a racecourse on which children of different ages and their elders can meet on an equal footing. Physical prowess and breadth of knowledge are immaterial on such a field. Only luck comes into play.

The playing of games has a long history. One can imagine the earliest humans engaged in contests of physical strength and endurance, with children racing about playing tag and great heroes struggling against daunting obstacles, as recorded in ancient myths. Written references to games go back thousands of years, and archaeologists have recovered a wide variety of relics that they interpret as gaming boards and pieces.

In the year 1283, when the king of Castile, Alfonso X (1221–1284), compiled the first book of games in European literature, he testified to the importance of games-playing in medieval society. "God has intended men to enjoy themselves with many games," he declared in the book's introduction. Such entertainments "bring them comfort and dispel their boredom."

Even in Alfonso's time, many of the board games he described were already hundreds of years old. Chess, the king's personal favorite, had been developed in India centuries earlier. Backgammon, one of the great entertainments of thirteenth-century nobility, had evolved from the Roman game *tabula*.

Succeeding centuries brought new amusements, along with variations on old ones. Each age and place had its particular favorites: the dice-and-counter game of pachisi in India, the coin-sliding game of shove ha'penny in William Shakespeare's England, the ancient game of go in China and Japan, and the card game cribbage in seventeenth-century Europe and America. In the Victorian era in Great Britain, nearly every parlor featured a wooden board of holes and pegs for the game of peg solitaire.

Amusement remains the motivation underlying the explosion of ingenuity that has now created a bewildering array of addictive computer, video, and arcade games, various forms of online and casino gambling, and new sports ranging from beach volleyball to snowboarding, along with novel board games and puzzles to tickle the mind.

"With their simple and unequivocal rules, [games] are like so many islands of order in the vague untidy chaos of experience," the novelist Aldous Huxley (1894–1963) wrote a few decades ago. "When we play games, or even when we watch them being played by others, we pass from the incomprehensible universe of given reality into a neat little man-made world, where everything is clear, purposive and easy to understand."

In these miniature worlds, competition brings excitement. Randomness serves as an equalizer. Chance introduces an element of suspense. Risk amplifies the thrill of play to an intoxicating level.

These tidy microcosms also attract mathematicians, who can't resist the distinctively human pleasure of learning the secrets of games. Who stands to win? What's the best move? Is there an optimal strategy? How long is a game likely to take? How do rules combine with chance to produce various outcomes? How are fairness and randomness linked?

In games of chance, each roll of a die, toss of a coin, turn of a card, or spin of a wheel brings a delicious surprise. Anyone can play. Anyone can win—or lose. Mathematics helps dispel some of the mystery surrounding unpredictable outcomes. It embodies an ever-present urge to tame the unruliness of Lady Fortune.

Using mathematical reasoning, we can't predict the outcome of a single roll of a die, but we can alter our expectations in the light of such an analysis. We may take comfort in the notion that, if the die is fair, each face of it will come up equally often in the long run. More generally, we can begin to make sense of and exploit the patterns that inevitably appear and disappear among the infinite possibilities offered by random choices.

Rolls and Flips

Dice are among the oldest known randomizers used in games of chance. In 49 B.C., when Julius Caesar ordered his troops across the river Rubicon to wage civil war in Italy, the *alea* of the well-known proverb he quoted already had the standard form of the die we use today: a cube engraved or painted with one to six dots, arranged so that the number of dots on opposite faces totals seven and the faces marked with one, two, and three dots go counterclockwise around a corner.

More than two thousand years earlier, the nobility of the Sumerian city of Ur in the Middle East played with tetrahedral dice. Carefully crafted from ivory or lapis lazuli, each die was marked on two of its four corners, and players presumably counted how many marked or unmarked tips faced upward when these dice were tossed. Egyptian tombs have yielded four-sided pencils of ivory and bone, which could be flung down or rolled to see which side faces uppermost. Cubic dice were used for games and gambling in classical Greece and in Iron Age settlements of northern Europe.

Because it has only two sides, a coin is the simplest kind of die. Typically, the two faces of a coin are made to look different (heads or tails), and this distinguishing feature plays a key role in innumerable pastimes in which a random decision hinges on the outcome of a coin toss.

How random is a coin toss? Using the equations of basic physics, it's possible to predict how long it takes a coin to drop from a known height. Apart from a small deviation due to measurement error, the time it will hit can be worked out precisely. On the other hand, a properly flipped coin tossed sufficiently high spins so often during its flight that calculating whether it lands heads or tails is practically impossible, even though the whole process is governed by well-defined physical laws.

Despite this unpredictability for individual flips, however, the results of coin tossing aren't haphazard. For a large number of tosses, the proportion of heads is very close to ½.

In the eighteenth century, the French naturalist Georges-Louis Leclerc (1707–1788), the Comte de Buffon, tested this notion by experiment. He tossed a coin 4,040 times, obtaining 2,048 heads (a proportion of 0.5069). Around 1900, the English mathematician Karl Pearson (1857–1936) persevered in tossing a coin twenty-four thousand times to get 12,012 heads (0.5005). During World War II, an English mathematician held as a prisoner of war in Germany passed the time in the same way, counting 5,067 heads in ten thousand coin tosses.

Such data suggest that a well-tossed fair coin is a satisfactory randomizer for achieving an equal balance between two possible outcomes. However, this equity of outcome doesn't necessarily apply to a coin that moves along the ground after a toss. An uneven distribution of mass between the two sides of the coin and the nature of its edge can bias the outcome to favor, say, tails over heads. A U.S. penny spinning on a surface rather than in the air, for example, comes up heads only 30 percent of the time. To ensure an equitable result, it's probably wise to catch a coin before it lands on some surface and rolls, spins, or bounces to a stop.

Empirical results from coin-tossing experiments support the logical assumption that each possible outcome of a coin toss has a probability of ½, or .5. Once we make this assumption, we can build abstract models that capture the probabilistic behavior of tossed coins—both the randomness of the individual tosses and the special kind of order that emerges from this process.

Consider what happens when a single coin is tossed repeatedly. On the first toss, the outcome is either a head or a tail. Two tosses have four (2×2) possible outcomes, each with a probability of ¼ (or .25); and three tosses have eight ($2 \times 2 \times 2$) possible outcomes. In general, the number of possible outcomes can be found by multiplying together as many 2s as there are tosses.

One can readily investigate the likelihood that certain patterns will appear in large numbers of consecutive tosses. For example, if a coin is tossed, say, 250 times, what's the longest run of consecutive heads that's likely to arise?

A simple argument gives us a rough estimate. Except on the first toss, a run of heads can begin only after a toss showing tails. Thus, because a tail is likely to come up about 125 times in 250 tosses, there are 125 opportunities to start a string of heads. For about half of these tails, the next toss will be a head. This gives us around sixty-three potential head runs. Roughly half the time, the first head will be followed by a second one. So, around thirty-two runs will consist of two heads or more. About half of these will contain at least one additional head, meaning that we will probably get sixteen runs of three heads or more, eight runs of at least four heads, four runs of at least five heads, two runs of six heads or more, and one run of seven heads or more.

That's actually a surprising result. People who are asked to write down a string of heads or tails that looks random rarely include sequences of more than four or five heads (or tails) in a row. In fact, it's generally quite easy to distinguish a human-generated sequence from a truly random sequence because the one that is written down by a human typically incorporates an insufficient number of long runs.

Possible Outcomes of Tossing
a Coin One, Two, or Three Times

	Number of Heads	Probability
One Toss		
T	0	½
H	1	½
Two Tosses		
TT	1	¼
TH	1	¼
HT	1	¼
HH	2	¼
Three Tosses		
TTT	0	⅛
TTH	1	⅛
THT	1	⅛
HTT	1	⅛
THH	2	⅛
HTH	2	⅛
HHT	2	⅛
HHH	3	⅛

In tossing a fair coin, the probability of each outcome—head (H) or tail (T)—is equal. If we toss once, there are only two possible outcomes, each of which has a probability of ½. Tossing twice, we have four possible outcomes, each having a probability of ¼. Tossing three times, we have eight possible outcomes, each having a probability of ⅛. From the table, you can see that with three tosses, the probability of obtaining no heads is ⅛, one head is ⅜, two heads is ⅜, and three heads is ⅛.

So, although an honest coin tends to come up heads about half the time, there's a good chance it will fall heads every time in a sequence of two, three, or four tosses. The chances of that happening ten times in a row are much smaller, but it can still happen. That's what makes it tricky to decide, just from a record of the outcomes of a short sequence of tosses, whether such as string is a chance occurrence or it represents evidence that the coin is biased to always come up heads.

Random Fairness

Like coins, cubic dice are subject to physical laws. An unscrupulous player can take advantage of this physics to manipulate chance. A cheat, for instance, can control a throw by spinning a die so that a particular face remains uppermost or by rolling it so that two faces stay vertical. In each case, the maneuver reduces the number of possible outcomes. A grossly oversized die, in particular, is quite vulnerable to such manipulation. The standardized size of dice used in casinos may very well represent a compromise configuration—based on long experience—that maximizes the opportunity for fairness. Casinos and gambling regulations specify the ideal dimensions and weight of dice.

A cheat can also doctor a die to increase the probability of or even guarantee certain outcomes. References to "loaded dice," in which one side is weighted so that a particular face falls uppermost, have been found in the literature of ancient Greece. Nowadays casino dice are transparent to reduce the chances of such a bias being introduced.

Even without deliberately creating a bias, it's difficult to manufacture dice accurately without introducing some asymmetry or nonuniformity. Manufacturers of casino dice take great pains to assure quality. Typically 0.75 inch wide, a die is precisely sawed from a rectangular rod of cellulose or some other transparent plastic. Pits are drilled about 0.017 inch deep into the faces of the cube, and the recesses are then filled in with paint of the same weight as the plastic that has been drilled out. The edges are generally sharp and square.

In contrast, ordinary store-bought dice, like those used in children's games, generally have recessed spots and distinctly rounded edges. Because much less care goes into the fabrication of such dice, they are probably somewhat biased. Achieving fairness is even more difficult with polyhedral dice that have eight, twelve, or twenty faces, each of which must be manufactured and finished to perfection.

In principle, an unbiased cubic die produces six possible outcomes. It makes sense to use a mathematical model in which each face has an equal probability of showing up. One can then calculate other probabilities, including how often a certain number is likely to come up. Several decades ago, the Harvard statistician Frederick Mosteller