have a kind of 999,999/1,000,000 proof. For instance, it is not uncommon for experts in DNA analysis to testify at a criminal trial that a DNA sample taken from a crime scene matches that taken from a suspect. How certain are such matches? When DNA evidence was first introduced, a number of experts testified that false positives are impossible in DNA testing. Today DNA experts regularly testify that the odds of a random person's matching the crime sample are less than 1 in 1 million or 1 in 1 billion. With those odds one could hardly blame a juror for thinking, *throw away the key*. But there is another statistic that is often not presented to the jury, one having to do with the fact that labs make errors, for instance, in collecting or handling a sample, by accidentally mixing or swapping samples, or by misinterpreting or incorrectly reporting results. Each of these errors is rare but not nearly as rare as a random match. The Philadelphia City Crime Laboratory, for instance, admitted that it had swapped the reference sample of the defendant and the victim in a rape case, and a testing firm called Cellmark Diagnostics admitted a similar error.[20] Unfortunately, the power of statistics relating to DNA presented in court is such that in Oklahoma a court sentenced a man named Timothy Durham to more than 3,100 years in prison even though eleven witnesses had placed him in another state at the time of the crime. It turned out that in the initial analysis the lab had failed to completely separate the DNA of the rapist and that of the victim in the fluid they tested, and the combination of the victim's and the rapist's DNA produced a positive result when compared with Durham's. A later retest turned up the error, and Durham was released after spending nearly four years in prison.[21]

Estimates of the error rate due to human causes vary, but many experts put it at around 1 percent. However, since the error rate of many labs has never been measured, courts often do not allow testimony on this overall statistic. Even if courts did allow testimony regarding false positives, how would jurors assess it? Most jurors assume that given the two types of error—the 1 in 1 billion accidental match and the 1 in 100 lab-error match—the overall error rate must be somewhere in between, say 1 in 500 million, which is still for

most jurors beyond a reasonable doubt. But employing the laws of probability, we find a much different answer.

The way to think of it is this: Since both errors are very unlikely, we can ignore the possibility that there is both an accidental match *and* a lab error. Therefore, we seek the probability that one error *or* the other occurred. That is given by our sum rule: it is the probability of a lab error (1 in 100) + the probability of an accidental match (1 in 1 billion). Since the latter is 10 million times smaller than the former, to a very good approximation the chance of both errors is the same as the chance of the more probable error—that is, the chances are 1 in 100. Given both possible causes, therefore, we should ignore the fancy expert testimony about the odds of accidental matches and focus instead on the much higher laboratory error rate—the very data courts often do not allow attorneys to present! And so the oft-repeated claims of DNA infallibility are exaggerated.

This is not an isolated issue. The use of mathematics in the modern legal system suffers from problems no less serious than those that arose in Rome so many centuries ago. One of the most famous cases illustrating the use and misuse of probability in law is *People v. Collins*, heard in 1968 by the California Supreme Court.[22] Here are the facts of the case as presented in the court decision:

> On June 18, 1964, about 11:30 a.m. Mrs. Juanita Brooks, who had been shopping, was walking home along an alley in the San Pedro area of the city of Los Angeles. She was pulling behind her a wicker basket carryall containing groceries and had her purse on top of the packages. She was using a cane. As she stooped down to pick up an empty carton, she was suddenly pushed to the ground by a person whom she neither saw nor heard approach. She was stunned by the fall and felt some pain. She managed to look up and saw a young woman running from the scene. According to Mrs. Brooks the latter appeared to weigh about 145 pounds, was wearing "something dark," and had hair "between a dark blond and a light blond," but lighter than the color of defendant Janet Collins' hair as it

> appeared at the trial. Immediately after the incident, Mrs. Brooks discovered that her purse, containing between $35 and $40, was missing.
>
> About the same time as the robbery, John Bass, who lived on the street at the end of the alley, was in front of his house watering his lawn. His attention was attracted by "a lot of crying and screaming" coming from the alley. As he looked in that direction, he saw a woman run out of the alley and enter a yellow automobile parked across the street from him. He was unable to give the make of the car. The car started off immediately and pulled wide around another parked vehicle so that in the narrow street it passed within six feet of Bass. The latter then saw that it was being driven by a male Negro, wearing a mustache and beard. . . . Other witnesses variously described the car as yellow, as yellow with an off-white top, and yellow with an egg-shell white top. The car was also described as being medium to large in size.

A few days after the incident a Los Angeles police officer spotted a yellow Lincoln with an off-white top in front of the defendants' home and spoke with them, explaining that he was investigating a robbery. He noted that the suspects fit the description of the man and woman who had committed the crime, except that the man did not have a beard, though he admitted that he sometimes wore one. Later that day the Los Angeles police arrested the two suspects, Malcolm Ricardo Collins, and his wife, Janet.

The evidence against the couple was scant, and the case rested heavily on the identification by the victim and the witness, John Bass. Unfortunately for the prosecution, neither proved to be a star on the witness stand. The victim could not identify Janet as the perpetrator and hadn't seen the driver at all. John Bass had not seen the perpetrator and said at the police lineup that he could not positively identify Malcolm Collins as the driver. And so, it seemed, the case was falling apart.

Enter the star witness, described in the California Supreme Court

opinion only as "an instructor of mathematics at a state college." This witness testified that the fact that the defendants were "a Caucasian woman with a blond ponytail . . . [and] a Negro with a beard and mustache" who drove a partly yellow automobile was enough to convict the couple. To illustrate its point, the prosecution presented this table, quoted here verbatim from the supreme court decision:

Characteristic	*Individual Probability*
Partly yellow automobile	1/10
Man with mustache	1/4
Negro man with beard	1/10
Girl with ponytail	1/10
Girl with blond hair	1/3
Interracial couple in car	1/1,000

The math instructor called by the prosecution said that the product rule applies to this data. By multiplying all the probabilities, one concludes that the chances of a couple fitting all these distinctive characteristics are 1 in 12 million. Accordingly, he said, one could infer that the chances that the couple was innocent were 1 in 12 million. The prosecutor then pointed out that these individual probabilities were estimates and invited the jurors to supply their own guesses and then do the math. He himself, he said, believed they were conservative estimates, and the probability he came up with employing the factors he assigned was more like 1 in 1 billion. The jury bought it and convicted the couple.

What is wrong with this picture? For one thing, as we've seen, in order to find a compound probability by multiplying the component probabilities, the categories have to be independent, and in this case they clearly aren't. For example, the table quotes the chance of observing a "Negro man with beard" as 1 in 10 and a "man with mustache" as 1 in 4. But most men with a beard also have a mustache, so if you observe a "Negro man with beard," the chances are no longer 1 in 4 that the man you observe has a mustache—they are much higher. That issue can be remedied if you eliminate the category "Negro man with beard." Then the product of the probabilities falls to about 1 in 1 million.

There is another error in the analysis: the relevant probability is not the one stated above—the probability that a couple selected at random will match the suspects' description. Rather, the relevant probability is the chance that a couple matching all these characteristics is the guilty couple. The former might be 1 in 1 million. But as for the latter, the population of the area adjoining the one where the crime was committed was several million, so you might reasonably expect there to be 2 or 3 couples in the area who matched the description. In that case the probability that a couple who matched the description was guilty, based on this evidence alone (which is pretty much all the prosecution had), is only 1 in 2 or 3. Hardly beyond a reasonable doubt. For these reasons the supreme court overturned Collins's conviction.

The use of probability and statistics in modern courtrooms is still a controversial subject. In the Collins case the California Supreme Court derided what it called "trial by mathematics," but it left the door open to more "proper applications of mathematical techniques." In the ensuing years, courts rarely considered mathematical arguments, but even when attorneys and judges don't quote explicit probabilities or mathematical theorems, they do often employ this sort of reasoning, as do jurors when they weigh the evidence. Moreover, statistical arguments are becoming increasingly important because of the necessity of assessing DNA evidence. Unfortunately, with this increased importance has not come increased understanding on the part of attorneys, judges, or juries. As explained by Thomas Lyon, who teaches probability and the law at the University of Southern California, "Few students take a probability in law course, and few attorneys feel it has a place."[23] In law as in other realms, the understanding of randomness can reveal hidden layers of truth, but only to those who possess the tools to uncover them. In the next chapter we shall consider the story of the first man to study those tools systematically.

CHAPTER 3

Finding Your Way through a Space of Possibilities

IN THE YEARS leading up to 1576, an oddly attired old man could be found roving with a strange, irregular gait up and down the streets of Rome, shouting occasionally to no one in particular and being listened to by no one at all. He had once been celebrated throughout Europe, a famous astrologer, physician to nobles of the court, chair of medicine at the University of Pavia. He had created enduring inventions, including a forerunner of the combination lock and the universal joint, which is used in automobiles today. He had published 131 books on a wide range of topics in philosophy, medicine, mathematics, and science. In 1576, however, he was a man with a past but no future, living in obscurity and abject poverty. In the late summer of that year he sat at his desk and wrote his final words, an ode to his favorite son, his oldest, who had been executed sixteen years earlier, at age twenty-six. The old man died on September 20, a few days shy of his seventy-fifth birthday. He had outlived two of his three children; at his death his surviving son was employed by the Inquisition as a professional torturer. That plum job was a reward for having given evidence against his father.

Before his death, Gerolamo Cardano burned 170 unpublished manuscripts.[1] Those sifting through his possessions found 111 that survived. One, written decades earlier and, from the looks of it, often