

# **BEAR IN THE MOONLIGHT**

*STORIES AND LESSONS IN  
PROBABILITY*

**BY BEN ORLIN**



## INTRODUCTION

Probability is the practice of quantifying uncertainty. It harnesses the power of mathematics to deal with our doubt, our ignorance, and our lack of guarantees in life. Like all of the best math, probability is brimming not just with practical applications, but with lovely ideas. It's like a gorgeous painting that also functions as a dishwasher.

Probability is beautiful, useful—and oh yeah, totally befuddling to most people who confront it. Consider some of the obstacles:

1. Probability is often **counterintuitive**. When dealing with uncertainties, our minds struggle to overcome built-in biases that run counter to logic and reason.
2. Probability is **not algorithmic**. In algebra, just memorizing the right steps will take you a long way (towards getting right answers, at least). But in probability, each question is a world unto itself.
3. Probabilities are expressed as **fractions**, which vex plenty of people in their own right.
4. Probability demands comfort with **very small** and **very large** quantities. For example, suppose I've got 20 different novels in a box. If I remove them randomly, one by one, what's the probability that they emerge in order from shortest to longest? Roughly 1 in 2 quintillion—a number beyond the typical limits of the human imagination.

5. Probability builds on **combinatorics**—the mathematics of sophisticated counting. Probability courses often begin with an intimidating unit on combinations, permutations, and the like. Conceptually, it's the right starting point. But pedagogically, it's awfully deep water for students just learning to swim.

Usually, someone learning probability tackles all these challenges at once. My hope is to isolate the first two obstacles: to help you wade into the non-algorithmic, counterintuitive nature of probability without getting drawn into the riptide of combinatorics and computations.

That's where the stories come in.

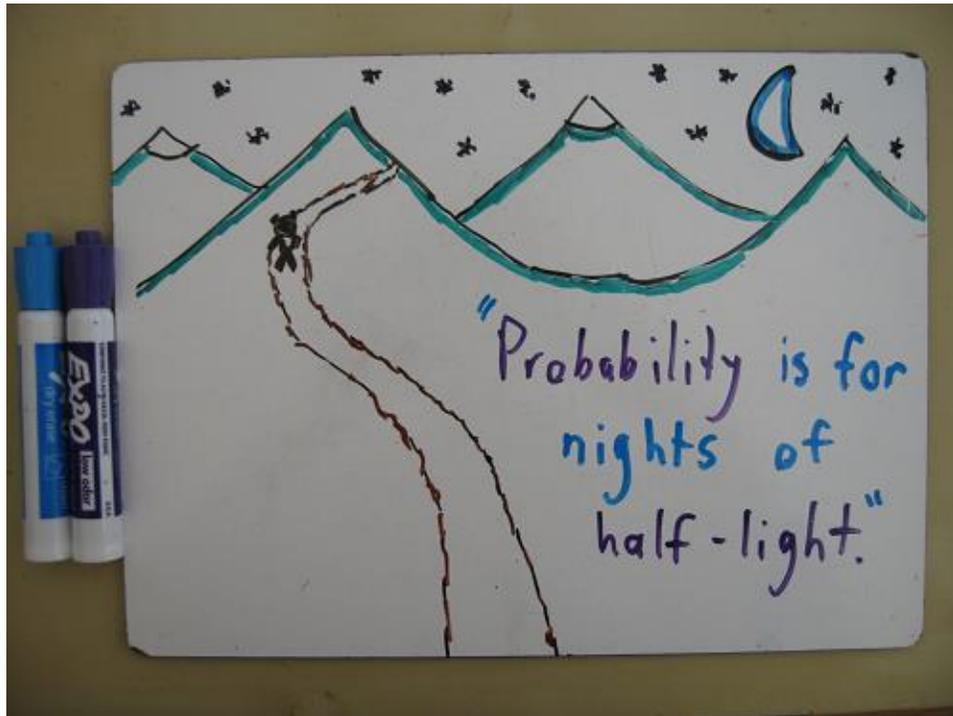
Narrative engages the mind. I've seen it happen in the classroom. A little story—even a clumsy or tangential one—grabs students in a way that few lectures do. Whereas concepts are so smooth that they slip right through our fingers, stories give us texture, a rough surface to grasp. Once engaged, we find our intuition and critical faculties (too often dormant in math class) hum to life. We're ready to wrestle with the big ideas, rather than crying "Uncle!" at the first sign of resistance.

Let's be clear. You won't learn probability just by reading stories. That'll take teachers, puzzles, struggles, and most of all, time. But I humbly offer these clumsy tales (with their even clumsier illustrations) in the hopes that they might spark a few insights or arguments.

After all, insights and arguments are what math is all about.



*These stories emerged from conversations with my father, James Orlin. We're currently working on a book about probability, which, unlike these stories, shall feature no monkeys. Sorry, monkey-and-math aficionados.*



## CH. 1: THE BEAR IN THE MOONLIGHT

*A probability is an expression of uncertainty.*

It was a half-moon that night. The student and the teacher could see a shadowy, white-chested figure lumbering down the mountain path.

"Is that a bear?" the student gasped.

The teacher nodded calmly. "It may be. Or, it may be one of the children from the village, disguised as a bear, hoping to scare his friends."

"Well, which is it?" the student hissed. "A deadly bear, or an innocent child?"

"Let us each determine the probability that the figure is a bear," the teacher said. "Then we shall share our answers with one another."

After a pause, the student whispered her answer. "20%. It could be a bear. But it looks too short, and I think it's wearing a backpack."

"Very good," the teacher said. "I say 40%. It moves slowly for a bear, but it seems to me the right size."

“So I’m wrong,” the student said. “It’s 40%.”

“No,” the teacher replied. “You are perfectly right. For me, it is 40%, and for you, 20%.”

“But you’re the teacher. You know more.”

“And your eyes are sharper than mine. Our perspectives are different, but neither is truer. I am right, and so are you.”

“So is it a bear,” the student said, with straining patience, “or not?”

The teacher closed her eyes. “What you seek is certainty. But a probability is only a perspective. Tell me, does that creature know whether it’s a bear or not?”

“Of course.”

“So for the creature itself, the probability must be 0% or 100%. It knows with certainty. You and I have our own perspectives, and thus our own probabilities.” The teacher paused. “Tell me, if there were a full moon tonight, what would we see?”

“It’d be bright,” the student said. “We could tell at a glance if that shadow is a bear.”

“And if it were a new moon, what would we see?”

“Nothing. Darkness. There would be no shadow at all.” The student paused. “We wouldn’t see the creature approaching, so we wouldn’t even be having this conversation.”

“Precisely. When the moon is full and bright, we know all. There is no need for probability. And when the moon is new and dark, we know nothing, not even enough to ask a question. In either case – total knowledge, or total ignorance – probability is useless.

“Probability is for the nights like these,” she continued. “It is for the nights of half-light. It is for the nights when we can make out a form, but cannot tell its precise shape. It is for nights when light and shadow mingle, when knowledge and ignorance share our thoughts. It is an expression of our uncertainty – no more, no less.”

“So you’re saying,” the student said, “a probability depends on what we know, and what we don’t know. And because you and I know different things, our probabilities are different.”

The teacher smiled. Looking back out the window, the student found that the figure—bear, child, whatever it was—had vanished.

## CH. 1: DISCUSSION

Probabilists list three approaches to probability: classical, empirical, and subjective.

“Classical” probability refers to chalkboard situations, like rolling dice and drawing cards, where we naively assume that a set of outcomes are equally likely. We take for granted, for example, that heads and tails are equally likely when flipping a coin—even though, in real life, a coin has microscopic imbalances, making one side slightly more likely than the other. Classical probability, then, is a purely theoretical game.

“Empirical” probability relies on real-world frequencies. For example, if 32% of skateboarders broke their noses while attempting a trick, we give you a 32% probability of the same.

“Subjective” probability aims to express uncertainty in our minds – and it’s much trickier to define. What does it mean to say that this shadow has a 20% chance of being a bear? There’s only one shadow. Or that the President has a 74% chance of winning reelection? There’s only one election. Or that tomorrow’s chance of rain is 30%? There’s only one tomorrow. Because we can’t repeat these events, it’s not obvious at first glance what such statements really mean.

## CH. 1: QUESTIONS

1. Which approach to probability—classical, empirical, or subjective—rings truest to your intuitions about what probability is?
2. Most probability teachers (myself included, honestly) prefer to sweep “subjective probability” under the rug. Why do you think that is?
3. Why, then, would the teacher in our fable want to introduce the student to probability via the subjective approach?
4. After the student hears the teacher say “40%,” should she adjust her own guess to something higher than 20%—like, say, 25% or 30%?
5. So the student says “20%,” then the teacher says “40%,” and then let’s suppose the student says, “I change my answer to 25%.” Should the *teacher* now change *her* answer? How long could this process go on? How long *should* it go on?
6. Suppose the student says the probability it’s a bear is 20%, and the probability it’s a child is 90%. What’s the problem with this? How does this illustrate a danger of subjective probability?



## CH. 2: THE BLINDFOLD AND THE CHESTNUTS

*It's all right to be blind. But don't pretend you can see.*

When the chestnuts finished roasting, a sweet aroma filled the kitchen. The student was about to dig in when, out of nowhere, a blindfold appeared in front of her eyes.

“No peeking,” the teacher warned.

The student heard the chestnuts being poured. “Now,” the teacher said, “I have divided your chestnuts among three bowls.

“I’ve also taken ten pieces of wasabi root, and carved them into the shape of chestnuts.” The teacher laughed. “These will not be to your liking. Bite into one, and your eyes will cry rivers, while your nose burns like a dragon’s.

“Now,” the teacher continued, “I am dividing my ten wasabi chestnuts among your delicious roasted ones. Six wasabi into the first bowl.” The student heard six plinking sounds. “Three into the second bowl.” More plinking sounds. “And one into the third.” Plink.

“You may reach into one bowl,” the teacher said, “and draw a chestnut at random. Which bowl do you choose?”

“The third, obviously,” the student said. “It only has one of your devil chestnuts.”

The student’s hand groped around the third bowl, but found only a single chestnut. Throwing off her blindfold, she saw that its color was a pale wasabi green. Peering into the bowls, this is what the student saw:

<b>Bowl</b>	<b>Regular Chestnuts</b>	<b>Wasabi “Chestnuts”</b>
First	100	6
Second	20	3
Third	0	1

“That’s not fair!” the student said. “You tricked me.”

“You tricked yourself,” the teacher said. “Why did you believe that the third bowl would be the best?”

“I figured the bowls would all have the same number of roasted chestnuts.”

“Why? Did anyone tell you this?”

“No,” the student said. “Once you said ‘three bowls,’ I just assumed you’d split the good chestnuts equally.”

“So what have you learned?”

“Well,” she said, “a probability is all about context. It doesn’t really matter how many wasabi nuts there are. It matters how many wasabi nuts there are *compared* with the other nuts. Even though the first bowl had the most wasabi, my probability of getting one was lowest, because there were so many other nuts, too.”

The teacher nodded. “What else?”

“You’ve got to know what information you’re missing, and what assumptions you’re making.”

“What else?”

“Never make a decision blindfolded.”

The teacher laughed. “An impossible wish. We’re all wearing blindfolds, every moment of our lives, and they come off far less easily than this cheap piece of cloth.”

“Then what should we do, when we can’t take the blindfold off?”

“Do the best you can,” the teacher said, “and never forget that you’re wearing it.”

## CH. 2: DISCUSSION

Lesson #1: A probability is a ratio. It's the number of outcomes you're interested in, divided by the total number of outcomes. A probabilist must remember that the numerator (in this case, the number of wasabi "devil chestnuts") isn't all that matters. You've also got to pay attention to the denominator (in this case, the total number of chestnuts).

Lesson #2: We often make hidden assumptions (in probability as in life). You can't avoid making assumptions altogether—for example, without *any* assumptions, the student could never have chosen a bowl—but it's important to know when you're making them. The assumptions that torment us most are the ones we're unaware of.

## CH. 2: QUESTIONS

1. Was the student wrong to reach for the third bowl? What would you have done?
2. More generally: What *should* a sensible person do in the student's place, and what does this tell us about how we should behave under uncertainty?
3. Give a way of dividing the 120 good chestnuts so that...
  - a. The third bowl would be the best choice
  - b. The second bowl would be the best choice
  - c. All three bowls would be equally good choices
4. Hidden assumptions are always dangerous in math. But why are they *especially* dangerous in solving probability problems?





## CH. 3: RIDDLE OF THE ODORLESS INCENSE

*See the world as it is, not as you imagine it to be.*

The teacher led the student to a blind vendor who sold two types of incense, identical in appearance. “This one he calls *Forest*,” the teacher said, holding up a speckled brown bundle that smelled of sandalwood and pine. “And this one he calls *Tea Garden*,” the teacher said, holding up another.

The student sniffed the *Tea Garden*. “It has no odor,” she said.

“Ah, a shame,” the teacher said. “Only some people can smell *Tea Garden*.”

The student shrugged. “Well, let’s not buy any of that one.”

“He sells them in bags of two,” the teacher continued. “But he does not pay attention to which incense goes in which bag.”

The blind vendor had a massive crate of incense, with the two types mixed together. His hands dove in and out, grabbing another stick with each motion, and throwing it into a plastic bag. When a bag contained two sticks, he set it aside for sale.

“I’ll just smell the bags before we buy them,” the student said. “And we’ll only buy ones where I detect *Forest*.”

“Ah,” the teacher said. “But won’t we end up paying for lots of *Tea Garden*, too?”

“The bags we’ll pick already have one *Forest*,” the student said. “So there’s a 50% chance they’ll contain another *Forest*, and a 50% chance they’ll contain a *Tea Garden*. That seems worth the risk.”

“Then let’s buy 120 bags,” the teacher said. “Start choosing.”

When they got home, the teacher put her to work immediately. “Smell each stick of incense, one by one,” she said. “Then make two piles of bags. The first is for those with two *Forest*. The second is for those with only one *Forest*.” The teacher smiled. “And we shall see which pile is larger.”

“They’ll be the same size,” the student said. “Like I said—there’s a 1 in 2 chance the second stick is *Forest*, and a 1 in 2 chance that it’s *Tea Garden*.”

“But how do you know which is the second stick?” The teacher giggled and walked off.

By the time she finished sorting, the student had grown convinced something was wrong. The first pile—with two-*Forest* bags—held only 41 bags. The other pile—the one-*Forest*, one-*Tea Garden* bags—stood almost twice as high, with 79 bags.

“The vendor cheated us,” the student seethed. “He deliberately gave me extra *Tea Garden*,” the student said. “Only half of our bags should have *Tea Garden*. But instead, 2 out of 3 do.”

They returned the next day to the vendor’s stall. “Don’t just watch him,” the teacher said. “Help fill the bags yourself. You must see the incense as the vendor does, not as the customer does.”

So the student began to fill bags, smelling each stick of incense as she grabbed it from the crate, and noting the order.

After half an hour, she exclaimed suddenly, “I get it! There are four types of bags.”

“Four?”

“Yes. There’s *Forest* plus *Forest*. And *Tea Garden* plus *Tea Garden*. And then, there are two more possibilities. There’s *Tea Garden* plus *Forest*, and there’s *Forest* plus *Tea Garden*.”

“Those last two,” the teacher said, “aren’t they the same?”

“The bags look the same when you’re done,” the student said, “but they’re not *created* the same. The process for each one is different.

“When we came yesterday,” the student continued, “we eliminated one of the four possible bags—the ones with only *Tea Garden*. That left three other types of bags. So of the ones we brought home, the double-*Forest* should only be 1 in 3. I thought I’d only picked bags where the first stick was *Forest*, but that was wrong. In some of the bags I picked, *Forest* was the second stick.”

When they arrived home, the student lit a stick of incense and sat with her eyes closed. “This *Forest* smells different than it did yesterday,” she said.

The teacher smiled. “I’m not surprised,” she said. “That’s *Tea Garden*.”

### CH. 3: DISCUSSION

This is a classic—and very tricky—problem. Usually, it’s presented the following way: “If a two-child family has at least one daughter, what is the probability that *both* their children are daughters?” Lots of people make the same mistake as the student in the story. The problem is an interesting introduction to the idea of “sample space,” because it requires careful thinking but almost no computation.

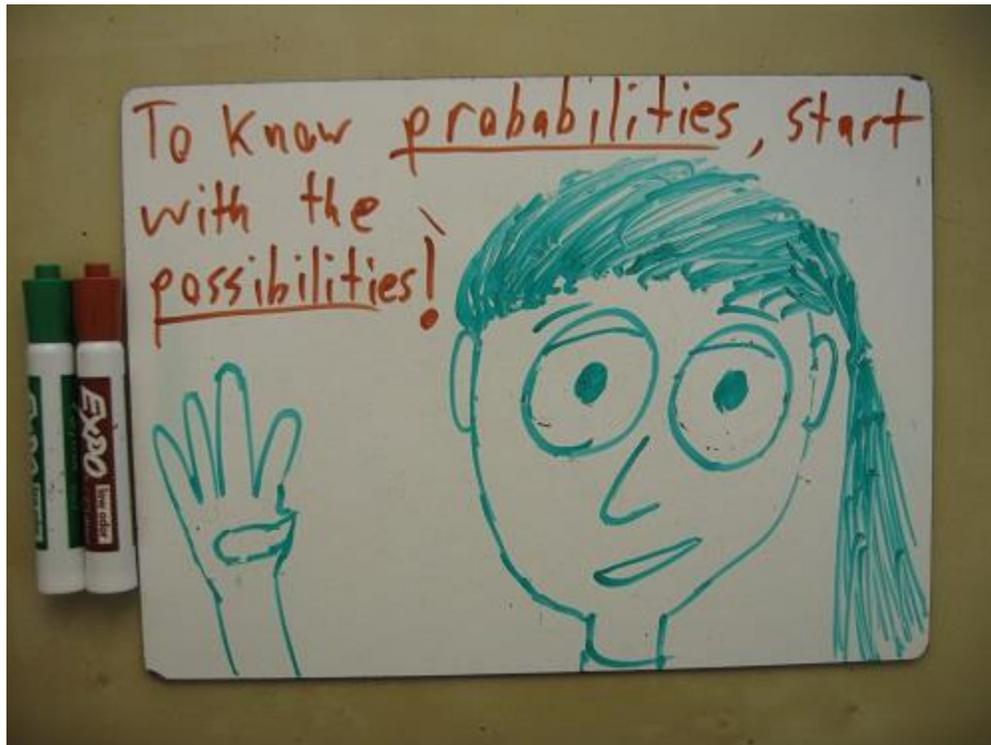
*Sample space*, by the way, is nothing but a fancy term for “list of possibilities.” The trick is that you’ve got to list possibilities that are equally likely. The student’s original sample space— $F + F, TG + TG$ , and “one of each”—was flawed, because the last item was twice as likely as either of the other two.

Technically, the incense scenario is slightly different from the daughter scenario. If you have a daughter, the probability your next child is also a daughter is precisely 50%. But if our crate of incense starts out with a 50-50 mix, then once we pick out a stick of *Forest*, there’s less *Forest* than *Tea Garden* remaining. So our probability is just below 50%. (Luckily, if there are thousands of sticks of incense, as in the story, then this change is barely noticeable.)

### CH. 3: QUESTIONS

1. Suppose the student picks a bag, then picks a random stick from inside and smells it alone. It’s *Forest*. What’s the probability that both sticks in the bag are *Forest*?
2. How is the scenario in #1 different from the scenario in the story? Does this matter?

3. *A case study in assumptions:* The student picks up a bag of two incense sticks. “Does this have any *Forest* in it?” she asks the vendor. He says yes. What is the probability that both sticks are *Forest*?
4. The vendor is now making sure that there are precisely two *Forest* sticks in every bag. But the bags now vary in size. Some have two sticks, but some have more. He just fills them, one stick at a time, until he reaches the second *Forest* stick. Then he moves on to the next bag. So by definition, the last stick he puts into each bag must be *Forest*. What’s the probability that the second-to-last stick in the bag is also *Forest*?
5. Suppose you’re picking sticks of incense, one after the other, until you obtain one of the following sequences: two straight *Forest*, or a *Tea Garden* followed by a *Forest*. When you hit one of those sequences, you stop picking. By definition, the last stick you pick must be *Forest*. But what’s the probability that the second-to-last stick is also *Forest*?





## CH. 4: THE SWINDLER'S COIN

*Everyone makes mistakes. Only fools stand by them.*

"I'm flipping a coin," the teacher said. "Tell me the probability that it comes up heads."

"Is it a fair coin?" the student asked.

"Yes," the teacher said. "I promise."

"Then the probability is 50%."

The student heard the *plink* of the flipped coin, and the slap as the teacher caught it. "Heads. Let's play again."

"50%," the student said. The coin spun through the air.

"Heads again," the teacher called. "Keep playing."

"50%," the student said. *A coin has two sides, so the probability of heads is 1 in 2. Easy,* she thought to herself.

"Heads again."

The conversation continued. “Another heads,” the teacher called. “We’re up to 30. What’s the probability that the next one is heads?”

“You’ve gotten really lucky, but the probability of heads is still 50%.”

The teacher shook her head. “You are a fool indeed.”

“But that’s how coins work,” the student said. “You told me it’s a fair coin, so the probability is 50%. All these heads have just been a wild coincidence.”

“Tell me,” the teacher said. “What is the probability of 30 heads in a row?”

“Um... 1 in a billion, more or less.”

“And there’s no chance that I’m swindling you, is there?”

“Well...” the student said. “It’s very unlikely. Not only would you have to lie to me, but you’d need to be able to manipulate coin tosses. Maybe there’s a 1 in a million chance.”

“1 in a million is quite different from impossible,” the teacher said.

“I guess.” The student shrugged.

“Now,” the teacher continued, “what is the probability that the next flip will be heads?”

The student paused. “I’m not sure. I thought it was 50%... but the chance of your tricking me is still 1000 times more likely than that you’re doing it honestly.”

“So is it 50%, or not?”

“It’s not,” the student concluded. “That’s a swindler’s coin.”

The teacher smiled. “Very good.”

The student’s face began to grow hot. “So you were lying to me this whole time! Why would you do that? You said it was a fair coin, and I trusted you.”

“As you should,” the teacher said, “up to a point. When you witness something that defies all of your assumptions about the world, you must learn to question those assumptions. To do otherwise is to disappear inside your own head, and ignore the world of evidence knocking at your door. You become one of three things: A dreamer, a fool, or a stubborn theoretician. And I’ve never had much luck telling the three apart.”

“You could have just told me that,” the student said. “You didn’t have to trick me.”

The teacher laughed. “How could I teach you about falsehoods, if I spoke only truths?”

## CH. 4: DISCUSSION RANT

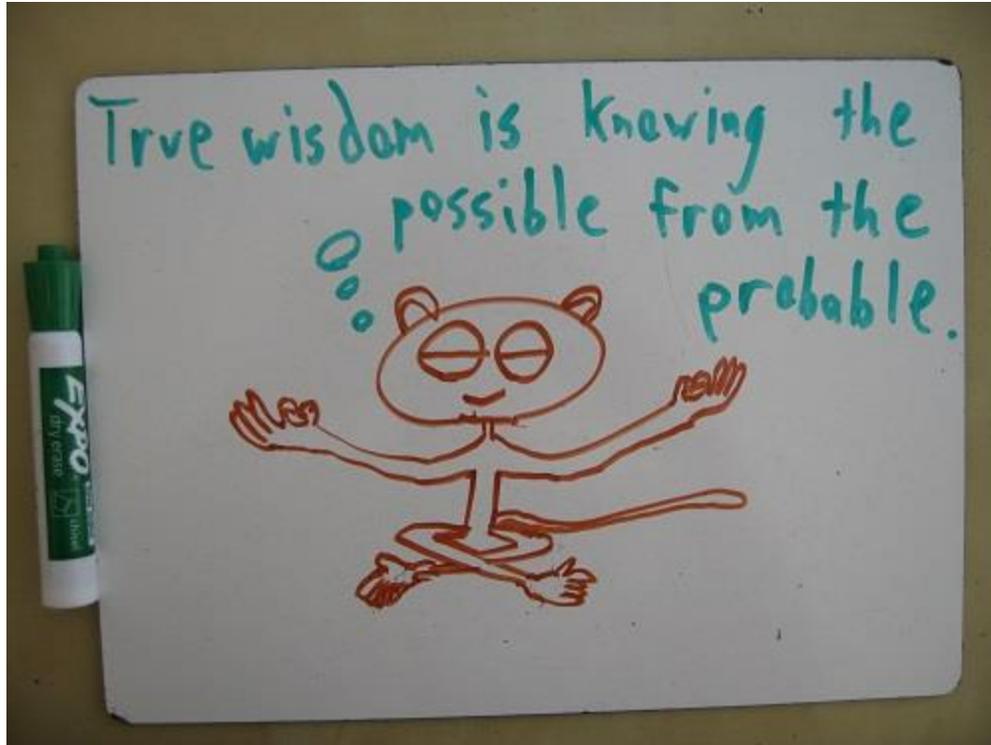
Some probability texts ask a similar question: "*If a fair coin is tossed 50 times, and comes up heads each time, what is the probability that it comes up heads on the 51<sup>st</sup> toss?*" The "correct" answer is  $\frac{1}{2}$ . A fair coin *always* has a probability  $\frac{1}{2}$  of coming up heads, because that's how we define "fair."

But guess what? If a coin comes up heads 50 times in a row—a 1-in-a-quadrillion event—then that ain't no fair coin. The question could be paraphrased: "If I tell you a coin is fair, and then overwhelming evidence accumulates to the contrary, would you still believe me?" And the "correct" answer would be: "Yes, because I never reconsider my assumptions."

For probability to be useful, it ought to stay anchored in practice. We shouldn't cling to invalidated assumptions or now-obsolete frameworks. We shouldn't keep telling ourselves the emperor is clothed.

## CH. 4: QUESTIONS

1. Suppose you're watching someone flip coins, and they keep getting heads. How many heads in a row would it take for you to believe that they're cheating somehow?
2. Suppose you're hearing somebody report the results of their coin flips (but you can't see for yourself). How many heads in a row would it take for you to believe they're lying?
3. Suppose you're flipping coins by yourself. How many heads in a row would it take for you to believe that someone is controlling the flips somehow?
4. Did you give the same answer to #1, #2, and #3? If not, what key factor(s) determine your willingness to believe in the coin's fairness?
5. What is Bayes' Law, and what does it have to do with this whole discussion?



## CH. 5: THE WISE MONKEY

*Read cautiously of rare events.*

The teacher had a new mission for the student. "Many monkeys live in the valley below," she said. "And it is said that in every 10,000 monkeys, there is a single one that possesses all knowledge."

"All knowledge?" the student said. "What, does it wear robes and talk?"

The teacher ignored her. "You ask it questions, and listen to its replies," the teacher said. "The Wise Monkey will coo for 'yes,' and grunt for 'no.'"

"And how am I supposed to find it?" the student said.

"Go to the valley," the teacher said. "Ask questions of the monkeys. If you find a Wise Monkey, it will answer correctly."

"And what if I find an ordinary monkey?"

"What do you think?" the teacher scoffed. "It will answer randomly."

The student set out for the valley. She moved from monkey to monkey, asking them questions until they got one wrong. A few monkeys showed promise, getting six, seven, or even

(in a few cases) eight questions right. But they all erred sooner or later. As sundown neared, the student began to despair.

“Is purple my favorite color?” she said, approaching yet another monkey.

The monkey grunted. No.

“Right,” the student said. “Is green my favorite color?”

The monkey cooed. Yes.

As the right answers mounted, the student could scarcely believe it. Ten questions later, the monkey was 10 for 10.

“I found one!” the student announced when she’d brought the monkey back with her.

The teacher studied her monkey carefully. “How do you know?”

“I asked it ten questions,” the student reported. “It got them all right.”

“Couldn’t that be a coincidence?”

“I *guess*,” the student said. “But it seems awfully unlikely. Ten is a lot of questions.”

“But tell me,” the teacher said, “how many monkeys did you speak with today?”

“Hundreds,” the student said. “Maybe a thousand.”

“I see.” The teacher nodded. “Let’s say a thousand. Now, of those, how many got the first question right?”

“Well, by dumb luck, half of them,” the student said. “So 500.”

“Good. And how many got the second question right, too – again by dumb luck?”

“Half of those. So 250.”

“And the third question?”

“Half again. 125.”

“And the fourth?”

“63, more or less.”

“And the fifth?”

“32.” The student got a sudden sinking feeling.

“And the sixth? Seventh? Eighth? Ninth?”

The student tried to keep up. “16. Then 8. Then 4. Then 2.”

“And tell me, my student,” the teacher said, “Of a thousand ordinary monkeys, how many would get all 10 questions right, purely by chance?”

The student slouched. “One would.”

The teacher then turned to the monkey. “Tell me, little one, is the sky blue by day?”

The monkey grunted. No.

“Really, now,” the teacher said. “Is the sky yellow by day?”

The monkey cooed, and the student groaned.

## CH. 5: DISCUSSION

This might remind you of the old adage that a thousand monkeys, given infinite time and a steady supply of typewriters, would reproduce the entire works of Shakespeare—minus the dubious “infinite time” part. Here, a thousand monkeys, given ten yes-or-no questions each, will produce one monkey who looks like a genius.

The moral amounts to this: Coincidences happen.

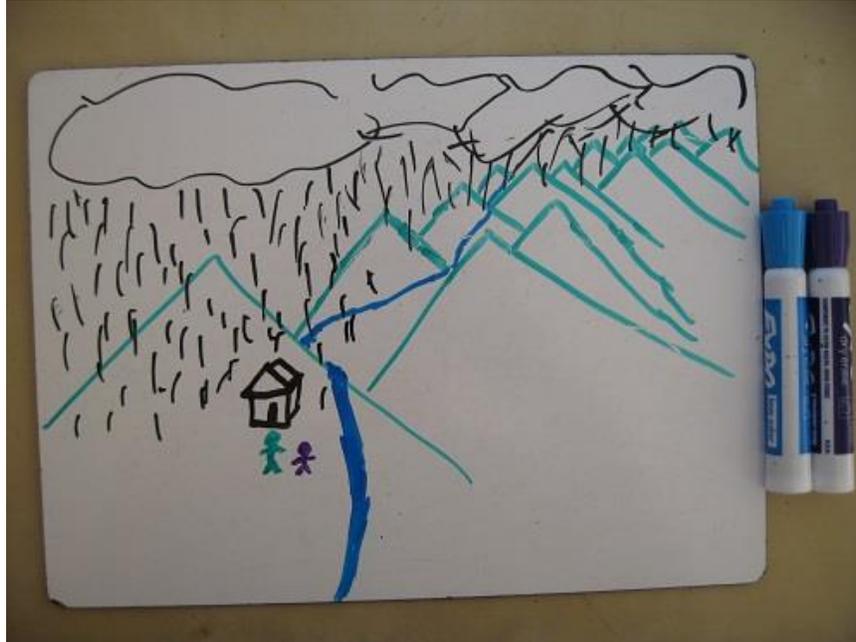
Think of sports—the octopus that accurately picked World Cup winners, or the pattern that the stock market dips when an AFC team wins the Super Bowl, or the countless fans who insist that their private rituals and practices steer their teams’ fortunes. You can chalk it all up to coincidence. If you try enough sea creatures, one is bound to predict the games correctly. If you look at enough meaningless indicators, one is bound to match the stock market’s fate, by sheer luck. And if you do enough random things each time your team plays, one of them is bound to correlate with the days your team wins.

But there are more pernicious examples. Consider money managers who claim they’re excellent at picking stocks. Gather together a thousand of these experts, and watch their performance for ten years. Track whether they perform above or below the market average. Even if their success is totally random, you’ll probably have someone who beats the market average for ten years running. Investors will flock to this wizard’s fund, but he might not be the Wise Monkey he seems to be.



## CH. 5: QUESTIONS

1. Let's say 1 in every 10,000 monkeys really *is* wise.
  - a. If you give a ten-question quiz to 10,000 monkeys, how many would you expect to get 100% on the quiz? (Include wise ones *and* lucky ones.)
  - b. Suppose we've got a monkey that aced the 10-question quiz. What's the probability that it's a wise one?
  - c. How many yes-or-no questions would a monkey need to get right before you'd be confident that it's a wise monkey, and not just a lucky imposter?
  
2. What if, instead of asking yes-or-no questions, the student asked the monkeys multiple-choice questions? (Assume the monkeys can make enough noises to answer these.) Out of 1000 (not-wise) monkeys, how many would get 100% on...
  - a. A five-question quiz with three answer choices per question (A, B, and C)?
  - b. A five-question quiz with four answer choices per question (A, B, C, and D)?
  - c. A five-question quiz with five answer choices per question (A, B, C, D, and E)?
  - d. Would a multiple-choice quiz have made it easier for the student to locate a wise monkey? Why or why not?
  
3. Here's a classic scam: Send out a letter to 32,000 people, promising to name a stock each week, and to state whether it will rise or drop in value. Give free tips for the first five weeks, and then charge \$1000 for the sixth week's tips. If you play your cards right, you could make up to \$1,000,000 that sixth week (minus the cost of postage). How does this scam work?
  
4. Why do humans often assign so much meaning to coincidences? Is this a bad impulse, a neutral one, or somehow beneficial?



## CH. 6: MOUNTAIN WHERE RAIN NEVER FALLS

*Use everything you know.*

Their sixth day of hiking brought the teacher and the student to an empty hut by a mountain stream. “We will rest here a while, and wash our clothes,” the teacher said.

When they had laid their clean clothes on sunny rocks to dry, the student pointed to the clouds gathering in the valley below. “Looks like rain. Should we be worried?”

“The rains have reached this place only once in the last 100 years,” the teacher said. “What is the probability that they will reach us today?”

The student thought for a moment. “Let’s say 30 storms reach the valley every year. Over 100 years, that’s 3,000 storms. But only one of them has reached the mountaintop here. So the probability is roughly 1 in 3000.”

“A very low probability,” the teacher nodded. Then she moved some of her half-dried clothes from the rocks to the inside of the hut. The student wondered why.

Later, the teacher plucked a blade of grass. “There is no dew,” she said. “This occurs less than once each month. What is the probability the rains reach us *now*?”

“Well, there’s no dew roughly 10 times per year,” the student said. “In the last century, that’s 1000 times. And at least 999 of those times, it didn’t rain. So the chances of rain are still at most 1 in 1000.” The teacher nodded, but moved more clothes inside nevertheless.

Later, a large brown bird flew directly overhead, barely 50 feet up. “The hawk-eagle flies low, just above the treetops,” the teacher said. “I glimpse such a sight only once a year. What is the probability the rain reaches us *now*?”

“Still small,” the student said. “The hawk-eagle flies low once a year, but it only rains once every century. So even if this is a sign of rain, the probability is still at most 1 in 100.” The teacher moved still more of her clothes inside the hut.

Later, as clouds began to darken the sky, the teacher pointed west, where pale stripes of color stood out against the gray sky. “A rainbow occurs in the West only once every 5 years,” the teacher said. “What is the probability that the rains reach us *now*?”

“Still low!” the student said. “There’s been a rainbow in the west 20 times in the last century. But it’s only rained one of those times. The probability is still just 1 in 20. There’s nothing to worry about.” Nevertheless, the teacher moved the last of her clothes inside the hut.

Just then, a light rain began to fall. The student scrambled to gather her clothes, but within minutes, the downpour was torrential, and everything was soaked through.

“I should have known better,” the student sighed from the shelter of the hut. “You tricked me again.”

“No trick,” the teacher said. “I told you everything.”

“How can that be?” the student said. “The odds should never have gotten higher than 1 in 20. What did you leave out? I used all three signs.”

“No.” The teacher shook her head. “You used one sign, then a different sign, then a third sign. You never used them all. How many days do you think *all three* signs have occurred?”

The student blinked. “I don’t know.”

“To my knowledge,” the teacher said, “only twice: today, and once many decades ago, when I was a young girl.”

“And you still remember that day?”

“Of course,” the teacher said. “It was the day that rain reached this place.”

## CH. 6: DISCUSSION

The focus here is *conditional* probability—the probability that one event will occur, given that another has already occurred.

In some sense, *all* probability is conditional. When we ask “What’s the chance of rain tomorrow?” we’re really asking, “What’s the chance of rain tomorrow, given the conditions today?” When we ask, “What’s the probability that the Patriots win the Super Bowl” we’re leaving out the rest of the question: “given their success so far this season?” To compute such probabilities, you’ve got to know what information to use, and how to use it.

The student in this parable deserves credit for applying the first key rule: *Finding out new information may change your probability.*

For example, say we’re rolling a standard die. What’s the probability we roll a 4? Simple: there are six possibilities, so it’s  $1/6$ .

But what’s the probability we roll a 4 *given that our roll is even*? Now we’ve got new information, and it changes our outlook. We can toss out 1, 3, and 5 as possibilities. That leaves 2, 4, and 6. So there’s a  $1/3$  chance that we rolled a 4.

Or, try this: What’s the probability we roll a 4 *given that our roll is a perfect square*? Well, there are only two perfect squares on a die—1 and 4. So the probability is  $1/2$ .

Finally: What’s the probability we roll a 4 given that our roll is even *and* a perfect square? There’s only one possibility fitting this description—4 itself. So the probability is 100%. It’s this final step that tripped up the student. She used one piece of information. Then she used the others. But she never used them all at once.



## CH. 6: QUESTIONS

1. Some of our idioms for “That’ll never happen!” can be reinterpreted as statements about conditional probability. Take your favorite low-probability event, and call it X. Then explain what each phrase has to do with conditional probability.
  - a. “If X happens, then pigs fly.”
  - b. “X happening? That’ll be the day hell freezes over.”
  
2. We say that A and B are *dependent* if knowing that one has occurred changes the probability of the other occurring. For example, when you flip one coin, getting heads and getting tails are *dependent*. Let’s see why.
  - a. What’s the probability that a coin flip comes up heads?
  - b. What’s the probability that a coin flip comes up heads, *given that it’s tails*? (Don’t overthink it.)
  - c. Why, then are “getting heads” and “getting tails” on a coin flip dependent?
  
3. Meanwhile, A and B are *independent* if knowing that one has occurred *doesn’t* change the probability of the other. For example, suppose we roll two dice, and we’re wondering two things: (A) Was the first die a 3? and (B) Was the sum of the dice 7? These might seem like dependent events, but they’re not. Let’s see why.
  - a. What’s the probability that the sum of two dice is 7?
  - b. What’s the probability that the sum of two dice is 7, *given that the first die is 3*?
  - c. Why, then, are “first die is a 3” and “sum of dice is 7” independent events?
  
4. Now, for each pair of events A and B, say whether they’re dependent or independent:
  - a. I pull one card from a deck. A = “the card is a diamond.” B = “the card is red.”
  - b. I pull one card from a deck. A = “the card is a diamond.” B = “the card is black.”
  - c. I pull two cards from a deck. A = “both cards are the same suit.” B = “the first card is a diamond.”
  - d. I pull two cards from a deck. A = “both cards are the same suit.” B = “both cards are the same number.”
  
5. Some people define “independent events” as “unrelated events.” It’s not a terrible shortcut, but it’s a little sloppy. Let’s see why.
  - a. Give an example of independent events that aren’t quite “unrelated.”
  - b. Give an example of “unrelated” events that might not be independent. (*Note: This obviously hinges a little on your definition of “unrelated.”*)



## CH. 7: PATTERNS IN THE STONEMWORK

*The eye sees chaos, and the mind imagines order.*

The teacher kept a garden, a grove of cypress and cedar with a pond at the center.

“You will build me a garden wall, alternating white stones and black,” the teacher told the student. “It must achieve the randomness and beauty of nature. I want to look at my wall and see a reflection of the cosmos.”

“How am I supposed to do that?” the student asked.

“Before you place each rock, flip a coin,” the teacher said. “If it comes up heads, place a white rock; if tails, place a black rock. That way, the sequence will be truly random, like nature.”

The teacher left, and the student began placing stones. But before long, flipping the coin grew tedious. Couldn’t she just pick a random color herself? The teacher would never know the difference. By the end of the day, she’d ringed half the garden with stones—white, black, black, white—and her fingers were cut and blistered.

The returning teacher glanced briefly at the wall. “The stones do not look random. You didn’t flip the coin.”

“How can you tell?” the student asked.

“The mind is full of false patterns,” the teacher said, “and this is what I see when I look at this wall. I do not see the random beauty of nature. I see the strained grunts and groans of a pattern-hunting mind.” The teacher shook her head. “Try the other half of the garden tomorrow. And this time, give me true randomness.”

When the student returned to work the following day, she obeyed the teacher’s instructions, and flipped a coin for each of the 500 stones. The teacher came home that evening and nodded. “Better work today.”

“What’s the difference?” the student asked.

“Look at your first day, the false randomness,” the teacher asked. “What is the longest streak of black stones that you lay, uninterrupted by white?”

“Four in a row.”

“Now look at the wall of true randomness.”

The student looked, and found long stretches of stones all the same color. One point had ten blacks in a row. By comparison, the first day of work looked quite regular, in its alternation between white and black.

“The truly random wall is full of streaks,” the student said. “But why is that?”

“Think,” the teacher said. “What is the probability of a streak of ten in a row?”

“About 1 in 500.” The student’s eyes widened. “And I laid 500 stones each day. So you’d *expect* to have a streak of ten somewhere in that sequence.”

“Yes,” the teacher said. “Your mind reads great meaning into streaks, and little meaning into their absence. But in true randomness, streaks are inevitable.”

“What I thought looked random, was actually full of patterns. The pattern was a lack of streaks,” the student said, pointing to the first day’s wall. “And what is actually random appears full of patterns, if you believe that streaks have meaning.”

The teacher nodded. Then they sat together for a while, admiring the clusters of cedar and cypress, and the gentle ripples in the water.

## **CH. 7: DISCUSSION**

This fable echoes Ch. 5, *The Wise Monkey*. Random processes (like flipping a coin) inevitably produce streaks and clusters that our minds interpret as meaningful patterns.

We struggle with randomness from both sides. As discussed earlier, we read significance into patterns that lack any interesting cause. Moreover, when we try to fake randomness, we hesitate to include coincidences and long streaks, so we create things that are too conspicuously even and balanced. So, paradoxically, in randomness we see patterns, and in patterns we see randomness.

This idea's applications extend beyond debunking false prognosticators—fortune telling, superstition, mutual funds, and the like. In its extreme version, this idea suggests that when we pour effort into explaining things, we're often wasting our time.

“Amazing—that NBA player just hit 8 three-pointers in a row! What's gotten into him?” Maybe nothing. Based on his career percentage you'd expect him to do that once or twice.

“This company's profits are soaring! What a genius CEO, right?” Well, with enough companies trying different strategies, some are bound to thrive by luck alone—not necessarily because of brilliant foresight.

“This school's test scores are phenomenal! What lessons can we draw?” Perhaps none—if talented teachers were distributed randomly across the country, a few lucky schools would receive a share far bigger than average, and we'd expect their students to excel.

We don't want to overextend this notion. It's not that causal explanations are never valid or valuable. But when you're seeking to explain some anomalous success or failure, it's worth asking—in a completely random world, wouldn't we expect a few outliers? Isn't it possible that this remarkable result comes—at least in part—from blind chance?



## CH. 7: QUESTIONS

1. Why do you think it's so hard for humans to generate random numbers or data? Why does everything we do—even when we're *trying* to be random—come out patterned?
2. Conversely: Why is it so hard for humans to accept “It's just random” as an explanation for patterns? Why do we find such non-explanations so unsatisfying?
3. Let's take the basketball example. (Choose a different sport if you prefer—it works out the same.)
  - a. Do you believe there's such a thing as “in the zone,” or “on a cold streak”? In other words, when playing a sport, are there times when you're more than just lucky or unlucky, but actually playing *better* or *worse* than average?
  - b. Let's suppose you answered “yes” to part (a). You think there *are* times when an athlete plays better or worse. (It seems plausible.) If that's the case, then when would you expect a basketball player to do better—right after he's missed his last two shots, or right after he's made his last two shots?
  - c. This gives us a testable prediction. If there's such a thing as “in the zone,” then a basketball player should have a better chance of scoring right after he's *made* two shots than after he's *missed* two shots. Take a guess: Do you think the data confirms this prediction?
4. An interesting case study: In the United States, the highest rates of kidney cancer (per capita in a town) occur in small, rural towns.
  - a. Come up with some plausible explanations for this fact.
  - b. Here's another fact: In the United States, the *lowest* rates of kidney *also* occur in small, rural towns. Try explaining *that*.
  - c. Finally, here's a hint: This is nothing more than randomness. What's going on?