THE KELLY CRITERION

The Wisdom of Betting

Seventy Years of the Kelly Formula

$$f^* = (b \cdot p - q) / b$$

Robert B.

F★ · 116 pp

For everyone who has ever sat staring at the screen of their account.

In 1956, in a small New Jersey town called Murray Hill, a thirty-two-year-old physicist wrote down half a line of text at Bell Labs. He probably never imagined that this half-line would go on, over the next seventy years, to save blackjack aces, the Bond King, Hong Kong racetrack punters, and Wall Street hedge fund managers — and to save you and me, too — so long as we were willing to listen to what it had to say.

This book is the story of that formula.

Contents

Preface · Half a Line on a Napkin

Many years ago, a young man sitting with me in a tavern slid a napkin across the table to me.

On it he had written a single line — a formula. It was so short that I felt a little let down. Only half a line long, with no integral sign, no Σ, none of the symbols that look "advanced":

b − q

$$f^* =$$

b

He said, "This is the Kelly formula. If you invest, you should do it by this."

I was young then, just starting out in investing, and I fancied I had something of a feel for mathematics. I glanced at that line and laughed:

"Isn't this something gamblers play with?"

My friend didn't argue with me. He folded the napkin twice and tucked it into the ashtray.

That was my first meeting with this formula. It was only many years later that I understood what I had so casually thrown away that night: a piece of wisdom that had already been tested by seventy years of time — one that had saved a blackjack ace's bankroll in the casino, saved the Bond King's two trillion dollars under management at PIMCO, saved the "professional syndicate" inside the Hong Kong Jockey Club, and saved the smartest crowd in the hedge fund world.

It also saved — could have saved — those who did not listen to it, and who in the end went to zero.

This book is not a textbook. There are enough textbooks already. Nor is it a memoir — I have no intention of telling you how much money I myself have lost.

This book is thirteen stories. In every one of them there is a real person, and his encounter with this formula.

• In the first story, there is a thirty-two-year-old physicist who, in the quiet corridors of Bell Labs, pushed a thought experiment — a gambler with an inside line to the phone — all the way to its limit. He died at only forty-one, collapsing on a Manhattan sidewalk. • In the second story, there is his colleague — Shannon, the man who invented "information theory." Shannon used this very formula to invest, earning 28% a year for thirty years. • In the fourth story, there is a math professor named Thorp who, in 1961, walked into a Reno casino with ten thousand dollars and a backer with mob connections, and in seventy-two hours walked away with $11,000 — bankrupting the myth that "commercial casinos are unbeatable." • In the eighth story, there is a physics master's graduate from Pittsburgh who drifted from Las Vegas to Hong Kong and, relying on a probability model with more than a hundred variables and a formula from 1956, made off with a billion dollars at the racetrack. • In the twelfth story, there are two Nobel laureates in economics who ran a hedge fund called LTCM at thirty-times leverage, and in three months in 1998 evaporated $4.6 billion — had they heeded this formula, they could have survived.

Between the stories lies hidden the seventy-year history of this formula's evolution:

• how it emerged from a paper that AT&T executives disliked enough to have renamed; • how it walked from the casino into Wall Street, and from Wall Street into Silicon Valley; • how it was humiliated by the Nobel laureate Samuelson in an essay written entirely in one-syllable English words; • how it survived — and how, here in 2026, it is being used by AI as "the last line of defense against humanity's urge to bet."

By the time you have finished these thirteen stories, you will discover something that surprised even me:

This formula, only half a line long, is at heart not about "how to win money," but about "how not to lose everything."

It has given all the truly great investors a shared restraint — whether it is the suit-wearing Bond King, or the horse-racing algorithm engineer writing code in Hong Kong's Central district, or that old man in Omaha sipping his Coke.

That napkin I threw away in the tavern all those years ago — it was only after many failures that I retrieved it from memory.

This book is the process of my retrieving it.

I hope it is of use to you.

Written in the spring of 2026

Part One · Birth

Chapter 1 · The Fighter-Pilot Physicist of Bell Labs

1.1 A Texas Boy

The day after Christmas in 1923, in Corsicana, Texas — an oil town — a boy was born. His father named him John Larry Kelly Jr., adding a "Junior" after the name in keeping with family tradition.

This child would one day change the history of finance. But there was no sign of it in his childhood.

Corsicana was an ordinary small town. Oil was its only industry. John did well in school from an early age; he loved mathematics, loved machinery, loved taking radios apart. He grew tall, with a Southerner's easygoing manner and courtesy.

When the Second World War broke out, he was in his teens. The moment he entered the Navy's flight school, he discovered that he liked airplanes more than he liked mathematics. He served four years as a fighter pilot.

When the war ended, he returned to the University of Texas at Austin to study physics. In 1953 he earned his doctorate in physics, with a dissertation in high-energy physics.

That year he was thirty.

1.2 Walking into Murray Hill

For a physics PhD in 1953, one of the most respectable destinations was an oil company. John first went into the oil industry — as a "research physicist," and he was paid well. But he soon found it boring.

In 1955, he moved to Bell Labs.

Bell Labs was the most magical place in the world at the time. It sat in Murray Hill, New Jersey — a tree-lined town so quiet it could lull you into a daze. It belonged to AT&T, the telephone monopoly of the entire United States. Out of the hundreds of millions of dollars in profit AT&T made each year, it siphoned off a portion to sustain this research institute.

This was where the transistor was invented (1947, Shockley), where the laser was invented (1958, Schawlow); this was where the Unix operating system and the C language were written; this was where information theory was born (1948, Shannon).

When John walked into Murray Hill, what excited him most was not the salary — it was the corridors. He could knock on any door at any time and chat with any Nobel-caliber colleague — this was Bell Labs' famous "open door" tradition.

He knocked on Shannon's door.

Claude Shannon was thirty-nine then, seven years older than John, and the "star" of Bell Labs. His 1948 paper, "A Mathematical Theory of Communication," was already treated as scripture by engineers the world over — it was the founding work of information theory.

In that paper Shannon proved one thing: information can be precisely quantified. There is a maximum rate at which a noisy channel can transmit information, called the "channel capacity," and it has a definite mathematical formula.

What was John doing at the time? No one remembers clearly. He may have been working on small communications projects, may have been doing odd jobs in signal processing. The center of gravity of his career had not yet surfaced.

But a few of his conversations with Shannon quietly set him thinking about a strange question —

1.3 A Thought Experiment: The Gambler with an "Inside Line"

The story begins somewhere between 1955 and 1956, when Bell Labs was doing some research on television signals.

By then Shannon was already mulling over a somewhat "philosophical" question: could the formulas of information theory (the channel capacity C) be applied outside communications engineering? To — the economy, say? To — investing?

Shannon and John had a few half-joking conversations in the corridors. Shannon's way of putting it went roughly like this:

"If you could know something a little earlier than everyone else — say, the result of a horse race — how much would that 'a little earlier' be worth?"

John was a Texan with no moral qualms whatsoever about gambling. He picked up the thread at once.

He imagined a gambler:

This gambler is placing bets. He has a private wire — a telephone line running straight to the back room of the racetrack. Before each race begins, the voice on the other end tells him the result. But this line is noisy — imperfect. Sometimes it tells him right, sometimes it tells him wrong.

Given the "signal-to-noise ratio" of this line, how should the gambler bet in order to make his wealth grow fastest over the long run?

This is the famous "private wire" thought experiment. It has two spine-chilling features:

First, it binds two things that look utterly unrelated — communications engineering (channel capacity) and finance (the rate of wealth growth) — to one and the same mathematical formula.

Second, the question it asks is not "how to gamble" — but "in the face of uncertainty, how much is information worth in return?"

1.4 The Formula Is Born

John spent a few months writing the thing down. The mathematics he used was not, in fact, complicated — roughly the stuff of a third-floor calculus course, which for a physics PhD was a warm-up exercise.

He first assumed the simplest case:

A wager has only two outcomes — win or lose. The probability of winning is p, the probability of losing is $q = 1 - p$. If you win, the odds are b (that is, betting 1 dollar wins you back b dollars, not counting the stake). You have a bankroll. Each time, what fraction f of your bankroll should you stake?

John's goal was not to maximize the single-round expectation — the strategy that maximizes single-round expectation is simple: bet everything. His goal was to maximize the long-run exponential growth rate of wealth.

This is an optimization problem a first-year calculus student could do. Write out the expected logarithm of the change in capital at each step:

$L(f) = p \cdot \log(1 + b \cdot f) + q \cdot \log(1 - f)$

Differentiate with respect to f, set the derivative equal to zero, and solve:

$$f^* = (b \cdot p - q) / b$$

This is the Kelly formula.

It tells you: when you have a positive expected edge, what fraction of your bankroll you should bet. This fraction is unique, definite, and objective.

Take an example. Suppose you flip a coin with a 55% win probability (you have a 5% edge), at odds of 1:1. Plug into the formula:

$$f^* = (1 \times 0.55 - 0.45) / 1 = 0.10$$

— bet 10% of your bankroll. Not 5% (too conservative), not 50% (too aggressive), and certainly not 100% (suicide) — just 10%.

This formula is beautiful in two respects:

First, its inputs are only two numbers — the win probability p and the odds b. No other nonsense. Second, its output is a fraction, not an amount — meaning that as your bankroll changes, your bet size adjusts automatically.

John also generalized the result to the multi-outcome case: a horse race with n horses, a wager with n different outcomes. He proved that under "no-arbitrage" odds (the house takes no cut), the optimal strategy is to allocate your money in proportion to the true probabilities — a property that would later play a crucial role in horse-racing algorithms (you will read Bill Benter's story in Chapter 8).

1.5 A Small Episode over the Paper's Title

John finished the paper and gave it a very straightforward title:

Information Theory and Gambling

He handed the manuscript to AT&T's legal and public-relations departments — by the rules of Bell Labs, every paper had to pass this review before publication.

The people in public relations saw the word "Gambling" and went pale.

"John," they said, "you cannot write 'gambling' on an AT&T paper."

AT&T was a telephone monopoly. It had moral responsibilities, a public image, and congressmen to placate. "A Bell Labs physicist is studying gambling" — if that made the newspapers, the anti-monopoly congressmen would tell it as a joke for three months.

John argued that he was not studying gambling; he was studying a financial interpretation of channel capacity.

The PR people said: explain it however you like, but the word "gambling" cannot appear in the title.

So he changed it. He changed it to this title:

A New Interpretation of Information Rate

This is a title that gives away nothing of the content. Any finance person who saw it would flip right past — it sounds like an engineering paper about the transmission rate of a telephone line.

In July 1956, the paper appeared in the Bell System Technical Journal, Volume 35, Issue 4, pages 917–926.

That journal was read mainly by telecommunications engineers. And so, on the day the Kelly formula was born, almost no one in the world of finance noticed it.

It lay quietly in an in-house journal of the Bell System, and waited five years.

1.6 An Interlude: The Man Who Made a Computer Sing

After John published the Kelly formula, his main line of research was, in fact, neither gambling nor finance.

He did one other thing at Bell Labs — a thing many people consider more famous than the Kelly formula.

In 1961, together with two colleagues, Carol Lochbaum and Lou Gerstman, he used an IBM 7090 computer to make the machine sing a song.

The song was called "Daisy Bell," an English love song from 1892. The computer sang the whole song in a synthesized "voice," accompanied by an arrangement made by another Bell Labs colleague, Max Mathews.

This was the first time in human history that a computer had been made to sing.

Word of this later reached a film director — Stanley Kubrick. While making "2001: A Space Odyssey" (released in 1968), Kubrick borrowed this scene: in the film, as the supercomputer HAL 9000 is slowly shut down, the last thing it sings is "Daisy Bell."

So — if you have seen "2001: A Space Odyssey" and remember that eerie, slow, warping "Daisy" as HAL dies — that is an echo of John's 1961 experiment at Bell Labs.

John, as a person, was more interesting than the formula he wrote. He liked doing things "that look useless." Making a computer sing has no practical use, but he did it in earnest.

This character trait may be precisely the reason the Kelly formula came out of him — he was not chasing usefulness; he was chasing the truth of a mathematical problem.

1.7 A Morning at Forty-One

March 18, 1965. Manhattan.

John was walking down the street that morning. He suddenly collapsed. A stroke. They took him to the hospital, but he could not be saved.

He was forty-one when he died.

He left behind three children. He had almost never stressed, in any of his papers, how important the Kelly formula was. He himself never once used the formula to gamble, nor used it to invest.

He may have been a man who was not good at explaining his own work to the outside world. Or he may simply not have cared about fame.

When he died, there was not a single financial institution in the world using the Kelly formula.

But —

One day in November 1960, five years earlier, Shannon had handed a photocopy of John's paper to a twenty-eight-year-old mathematics instructor at MIT, and said to him:

"You should read this."

That mathematics instructor was named Edward O. Thorp.

John's story ended within his own timeline. But his formula was only just beginning its journey.

1.8 What This Formula Is Really "Saying"

Let us pause the story and return to the formula:

$$f^* = (b \cdot p - q) / b$$

The literal meaning of this formula is "what fraction of your bankroll to bet." But what it is really saying goes far deeper than "bet size."

It is saying: the size of your bet should be proportional to your "informational edge."

• The greater the edge (the larger $bp - q$), the more you should bet. • When the edge is zero ($bp - q = 0$), you should not bet. • When the edge is negative ($bp - q < 0$), you should bet the other way or step out.

It is also saying: your bet should scale in step with your wealth — always "some fraction of your current bankroll," never a fixed amount. This is why the Kelly formula is naturally suited to compounding: when you win, your next bet size automatically grows larger; when you lose, your next bet size automatically shrinks.

These two properties together make the Kelly formula the strategy that maximizes the rate of wealth growth in a long-run game.

In the decades that followed, mathematicians would give more rigorous proofs (you will read Breiman's story in Chapter 3), a Nobel laureate would leap out in opposition (you will read Samuelson's story in Chapter 11), there would be people who used it to beat the casino (Chapters 4–5), people who used it to beat the market (Chapters 6–7), people who used it to beat the racetrack (Chapter 8), and people who ignored it entirely and died (Chapter 12).

But the source of all these stories is that one paper with its retitled name, at Murray Hill in 1956, and the half a line that a thirty-two-year-old fighter-pilot physicist wrote down on his scratch paper.

Chapter Summary: The Three Things John Kelly Left Behind

The first: a formula only half a line long, $f^* = (b \cdot p - q) / b$, telling you, in a wager with known win probability and odds, what fraction of your bankroll you should bet.

The second: a profound mathematical equation — channel capacity = long-run rate of wealth growth. In other words, information can be translated one-to-one into return. One more bit of information, one more measure of growth.

The third: a surprising fact — the inventor of this formula never once used it to invest. All of its later applications in the world of finance were carried out by others.

When John Kelly died in 1965, he was forty-one. He did not live to see the day the Kelly formula beat the casino, beat Wall Street, beat the racetrack.

But his colleague and friend — Claude Shannon — had already handed this formula, on his behalf, to a man who really would put it to use.

That was 1960. The man was twenty-eight. His name was Edward Thorp.

See you in the next chapter.

Chapter 2 · Shannon's "Instigation"

2.1 A Letter Arrives in Cambridge

November 1960. Cambridge, Massachusetts. The Mathematics Department at MIT.

A twenty-eight-year-old instructor opened a letter. It had come from an unexpected place—Bell Labs, Murray Hill, New Jersey. The sender was Claude Shannon.

The instructor was named Edward O. Thorp. He had been at MIT for only a year. What he was doing struck his colleagues in the math department as a little . . . odd.

He was using an IBM 704 computer to study blackjack.

Not because he loved gambling—he didn't. It was because he believed one thing: that a casino game like blackjack could, mathematically, be beaten. If he could prove it, he would rewrite the history of gambling.

But to publish a paper, he needed a member of the National Academy of Sciences (NAS) to forward it and put his name to a recommendation—that was the rule at the Proceedings of the National Academy of Sciences. Among the academicians in MIT's math department, no one was willing to vouch for it: a mathematics PhD dabbling in gambling? It didn't sound serious.

Thorp wouldn't give up. He asked around: who among the mathematically inclined cared least about whether something was "serious" or not?

The answer was: Shannon.

By then Shannon was already teaching at MIT. He had been "seconded" from Bell Labs to MIT, but he still shuttled between the two. He was famous in an uncanny way—a fellow who would ride a unicycle down the hallway, whose office was piled with homemade juggling contraptions, a homemade chess-playing machine, a homemade mechanical arm that could solve a Rubik's Cube.

Thorp arranged a meeting. The secretary told him: "Professor Shannon can give you only five minutes—he doesn't like to be disturbed."

Thorp walked in, and his very first sentence was:

"I've used mathematics to beat blackjack."

Shannon's eyes lit up.

Five minutes turned into five hours.

2.2 The Man Called Shannon

To understand why Shannon so readily seized on Thorp's topic, you first have to understand the man himself.

He was born in 1916, seven years older than Kelly. From childhood he possessed a temperament that fused a child's curiosity with an engineer's knack for building things. While earning his master's at MIT, he wrote a thesis: using Boolean algebra to describe electrical circuits. That sounds like a small subject, but in fact—it laid the logical foundation of every modern computer.

Howard Gardner, the Harvard scholar, later said that this 1937 master's thesis was "possibly the most important master's thesis of the twentieth century, bar none."

At twenty-two, Shannon laid the logical foundation of the computer. At thirty-two (in 1948), he wrote "A Mathematical Theory of Communication"—the founding text of information theory.

In other words, one man, twice, rewrote the twentieth century.

But he never showed off. At Bell Labs and at MIT, what made him famous was not his papers, but his toys:

• He built an "Ultimate Machine": you flip the switch, the machine's lid opens, a mechanical hand reaches out, turns the switch off, then retracts, and the lid closes. The machine's sole purpose in existing was to turn itself off. • He built a "mechanical mouse" named Theseus that could find its own way through a maze, and on the second run would go straight down the shortest path—an early prototype of artificial intelligence. • He rode a unicycle, and while riding it he juggled three balls. He even wrote a paper on "the mathematics of juggling," proving a "juggling theorem." • In 1949—twenty years before IBM—he wrote a chess-playing program on components he had built himself.

Shannon was nothing like a finance man, still less like a gambler. He was like an engineering student who never quite grew up.

But it was precisely this "want-to-try-everything" curiosity that made him, when he met Thorp in 1960, react instantly:

"Blackjack? Beating the casino with mathematics? Isn't that just an extension of that whole Kelly-formula business? Come on, let's give it a try."

2.3 The Two Gifts Shannon Gave Thorp That Day

Thorp said later, in his memoirs, that on that day in 1960 Shannon gave him two gifts—ones he remembered his whole life.

Gift one: the signature on that NAS paper. Shannon agreed on the spot, and within a few days he helped Thorp forward, recommend, and publish the paper. In January 1961, Thorp's "winning strategy for blackjack" was formally published in the Proceedings of the National Academy of Sciences. For the first time, academia acknowledged: the casino is not invincible.

Gift two: a photocopy of another paper. Shannon pulled a sheaf of pages from a drawer and handed it to Thorp:

"You ought to read this."

It was the paper Kelly had published in 1956 in the Bell System Technical Journal—A New Interpretation of Information Rate.

Shannon said: "This was written by my old colleague at Bell Labs, John Kelly. He used my information theory to derive a formula that tells you how much you should bet. If you're going to play blackjack for real at the table, you have to understand this first."

Thorp took the sheaf of pages home.

He said later that he read Kelly's paper three times that night. The next day he read it three times more.

He grasped one thing: he knew how to compute the win probability (blackjack's "card-counting method"), but he didn't know how much to bet on each hand. The Kelly formula was precisely what filled that gap.

"Blackjack's card counting tells you how much edge you have; the Kelly formula tells you how much to bet. Only the two together make a complete weapon."

Thorp went on to publish Beat the Dealer in 1962. That book became a legend in the history of American publishing—the first bestseller to teach ordinary people how to beat the casino with mathematics. In it, Thorp gave the formula, for the first time, a formal name—

"the Kelly gambling system."

Later, in the 1966 revised edition, he fixed the name as:

"the Kelly Criterion."

From that moment on, John Kelly's 1956 paper, which no one had read, had its name. That name would accompany it through the next seventy years.

2.4 A Subplot That Cannot Be Written: Shannon and Thorp's Las Vegas

Having come this far, the story actually has a subplot that cannot go untold —

Thorp and Shannon, in the summer of 1961, really did go to Las Vegas together.

And what they brought with them was more than the Kelly formula. What they brought was the first wearable computer in human history.

Here is how it happened:

Beyond blackjack, Thorp also wanted to crack roulette. Roulette looks completely random—the little ball rolls around the wheel, and you guess which number it will land on. But Thorp asked an engineer's kind of question: "If I measure the speed of the ball and the speed of the wheel, can I predict, before the ball settles, which half it will land in?"

He and Shannon, in the evenings down in the basement of Shannon's house—a room piled with mechanical parts, soldering irons, and wires—assembled together a small computer about the size of a shoebox. This computer could be tucked into a trouser leg, hidden in the sole of a shoe, operated with the toes on a set of switches, with the result heard through a miniature earpiece.

The procedure:

1. Thorp, wearing this computer, would walk into the casino and stand by the roulette wheel.
2. The wheel would spin, and the ball would be thrown in.
3. Thorp would press a switch with his toe, telling the computer "the moment the ball first passes number 12"—from which the computer would estimate the ball's speed.
4. Another press with the toe, telling the computer "the moment the wheel passes the reference point for the second time"—from which the computer would estimate the wheel's speed.
5. The computer, using a simple physical model, would predict roughly which one-eighth arc of the circle the ball would land in.

6. Through a miniature earpiece, it would tell Thorp the answer using one of eight tones.
7. Thorp would hurry to place his chips on the corresponding eight numbers.

This was the first wearable computer in human history, fifty years ahead of the so-called "smartwatch."

Its effect in practice: a theoretical edge of about 44%. But there was one engineering problem—the earpiece wires kept breaking. Thorp and Shannon fixed them again and again, and again and again their runs in the casino were ruined by the earpiece failing. They didn't win much big money (nor lose much), but the theory was borne out.

Shannon later donated this prototype to the MIT Museum. It is still on display there today.

Meanwhile, on the blackjack side, Thorp walked into a Reno casino with another backer—Manny Kimmel. Kimmel was a wealthy man with mob ties, who had once been a bootlegger, run underground lotteries, and had dealings with New York crime syndicates. But he had money, he had nerve, and he was willing to put up $10,000 (about $110,000 in 2026 dollars) to let Thorp play for real.

Seventy-two hours later, they had netted $11,000—more than doubling the stake.

This was the first real-world profit from the Kelly formula. We'll save that story for a fuller telling in Chapter 4.

2.5 Shannon Himself Used the Kelly Formula Too

One thing many people don't know is this: Shannon himself was also an investor, and a very successful one.

He began investing in American stocks in the 1950s. He almost never traded—the stocks he bought he simply clung to, and did not sell.

His core position was three stocks:

• Teledyne (an industrial conglomerate) • Motorola • Hewlett-Packard

He bought into these three stocks very early. Teledyne he invested in when the company was still very small—he knew the founder and was bullish on their engineering team. That stock went on to become a legend of the 1960s—with an astonishing annual compound growth rate.

By Shannon's own estimate, his annualized compound return over thirty years (roughly 1960–1990) was about 28%. What does that number mean? Over the same period, the S&P 500's annualized return was about 10%. A 28% annual compounding sustained over thirty years equals a 1,645-fold multiplication of capital.

He followed three principles in his investing:

1. Look at a company's fundamentals, not at the fluctuations of its stock price. "A lot of people stare at the stock price; they ought to be watching the company's earnings."
2. Choosing the companies that will succeed is far easier than predicting short-term fluctuations.
3. Bet cautiously—he never went all-in on a single stock; he always kept cash on hand, leaving himself the room to "act when the opportunity comes."

The third principle is a very characteristic piece of the Kelly spirit.

But Shannon went further. He also proposed a thought experiment called "Shannon's Demon": by continually rebalancing between a volatile stock and cash, you can create a portfolio that grows over the long run—even if the stock's own geometric mean is negative!

This sounds like financial alchemy, but Shannon rigorously proved that it holds under certain conditions. This "demon" later became a classic concept in the world of quantitative investing, called "volatility harvesting."

2.6 Kelly and Shannon: A Pair of Well-Matched Collaborators

By this point you may want to ask: between Kelly and Shannon, who was more important?

My view is: it is the two of them together that make up the Kelly formula.

Kelly supplied the financial insight—turning information theory into a betting proportion, translating "channel capacity" into "rate of wealth growth." The key to this was not mathematical difficulty, but the courage to cross intellectual boundaries. To get a thirty-year-old physics PhD to bind communications engineering and gambling together—what that required was not technique, but a mind not fenced in by "disciplinary boundaries."

Shannon supplied the theoretical foundation and the network of transmission. His information theory was the mathematical womb of the Kelly formula. And the human bridge he built between MIT and Bell Labs allowed this formula to walk out of Kelly's drawer—passed into Thorp's hands, and into the hands of one generation of quantitative investors after another.

At a deeper level: both Kelly and Shannon were people who "did not do research along disciplinary lines."

Kelly was a physicist, yet he wrote a finance paper. Shannon was an electrical engineer, yet he wrote a foundational paper in probability theory. Both of them carried a temperament peculiar to Bell Labs:

Your training background doesn't matter. What matters is where the interesting problems are.

This temperament is rarely seen in today's academy. Today's academic system nails everyone into his own little box—if you do physics, you can only submit to physics journals; if you do finance, you can only use the terminology of finance. The most interesting insights often die on these very boundaries.

The Kelly formula is an exception. It grew out from the information-theory side and reached over to the finance side—and this cross-boundary tree bore seventy years of fruit.

2.7 Shannon's Old Age and Kelly's Early Death

The story must have a contrast.

Kelly died of a stroke in 1965, at forty-one. He did not see his formula beat the casino, did not see Beat the Dealer published (in fact it was published in 1962, and he may have heard of it, but he absolutely did not see the 1966 revised edition write his name into the term "the Kelly Criterion"). He did not live to see Thorp open Princeton-Newport Partners on Wall Street in the

1970s, nor to see Bill Benter in Hong Kong in the 1980s use his formula to make a billion.

Like so many geniuses of that Bell Labs era, he died early and was famous late.

Shannon lived to eighty-four. But in his later years he developed Alzheimer's disease—the man who proposed "information theory," in the last ten years of his life, slowly lost all his information. In February 2001, he died in a nursing home in Massachusetts.

By the time he died, the Kelly formula had already become a standard tool in the hedge fund world. Thorp had already closed PNP (in 1988), and had later opened Ridgeline Partners (a 21% net annualized return, 1994–2002). Bill Benter had already amassed a fortune of several hundred million dollars in Hong Kong. Bill Gross was managing trillions of dollars in bonds at PIMCO.

Shannon probably knew about all of this—at least while he was still lucid.

But he never regarded the Kelly formula as "my invention." He always called it "John Kelly's formula."

This is one of the very costly qualities of the man Shannon—he never credited another's work to himself.

Chapter Summary: The Three Things Shannon Left to Posterity

The first: he passed Kelly's paper into Thorp's hands. This one act was more important than if he had directly invented the Kelly formula himself—because it was this act that carried the Kelly formula out of Bell Labs.

The second: with his own hard cash he proved that the spirit of the Kelly formula works—28% annualized over thirty years, holding a handful of quality companies for the long run, and always keeping cash on hand to meet opportunity.

The third: he left posterity a Bell Labs–style attitude toward research—not to be confined by disciplinary boundaries, and to step across and take a look whenever an interesting problem appears. In today's age of artificial intelligence, this attitude may be even scarcer than "information theory" itself.

In the next chapter, we will look at two other mathematicians who made the Kelly formula truly "stand on its feet"—one named Henry Latané, the other named Leo Breiman. What they did was not as dramatic as Thorp's story, but they are the reason this formula truly "stands up mathematically."

Without them, the Kelly formula might to this day still be nothing more than "a clever little trick from an oddball at Bell Labs," rather than a cornerstone of financial mathematics.

Chapter 3 · The Mathematicians Who Gave the Formula a Soul

3.1 A Forgotten Financial Economist

1959. The University of North Carolina at Chapel Hill.

A professor of financial economics named Henry Allen Latané mailed a paper to the Journal of Political Economy.

This journal is one of the world's top economics periodicals, published by the University of Chicago; within the field of economics its standing is second only to the American Economic Review. Latané's paper was not long, and its title was plain:

Criteria for Choice Among Risky Ventures

In the article he discussed something that sounds like a philosophical question:

Suppose you have several mutually exclusive investment opportunities, each with a different expected return and risk. Which one should you choose?

The traditional answer was the mean-variance analysis proposed by Harry Markowitz in 1952 — find a portfolio that gives you "the maximum expected value for a given variance" or "the minimum variance for a given expected value."

But Latané disagreed. He said: for someone who will invest repeatedly, many times over, the correct criterion is not "mean-variance" but "geometric mean maximization."

Specifically, you should choose the option that maximizes the geometric mean of your wealth — in the long run, the option with the highest geometric mean will grow your wealth the fastest.

This sounds like exactly what the Kelly formula says — because geometric mean maximization and logarithmic utility maximization are equivalent, and the solution to logarithmic utility maximization is precisely the Kelly formula.

But Latané's work was done independently. He had not read Kelly's 1956 paper published in the Bell System Technical Journal — why would a financial economist read a telecommunications engineering journal?

Two men, from two entirely different fields, independently reached the same conclusion.

Such "independent discovery" is not rare in the history of science — it usually means the conclusion is correct, because it does not depend on any particular path of research.

3.2 Latané's Contribution: Bringing Kelly into Financial Economics

Latané's paper did several things that truly got the Kelly formula "in the door" of finance:

First, in the language of a financial economist, he made "geometric mean maximization" clear as an investment criterion. Kelly's paper was written for telecommunications engineers, full of terms like channel capacity and signal-to-noise ratio, terms utterly unfamiliar to finance people. Latané translated it into words that finance people could understand.

Second, he showed that the geometric mean criterion has a very benign statistical property — decreasing absolute risk aversion. This term sounds very academic, but its meaning is actually quite intuitive: the wealthier you are, the lower your aversion to a loss of the same amount should be. This is the true structure of preferences for most rational investors.

Third, he got Markowitz's attention. In Chapter Six of his 1959 book Portfolio Selection, Markowitz specifically discussed geometric mean maximization as a preference basis for long-term investing. Markowitz later won the Nobel Prize in Economics in 1990 — his middle-ground stance toward the Kelly direction remains, to this day, the most authoritative "non-opposition" voice on the Kelly formula within the field of finance.

Latané himself never became a star of the finance world. He spent his whole life as a professor at the University of North Carolina, wrote a few not particularly famous finance textbooks, retired in 1984, and died in 1996. He almost never publicly discussed the Kelly formula — he did not know that the 1959 paper he had written would, forty years later, become one of the foundational documents of quantitative finance.

This is a common thing in the history of scholarship: some people do important work but never realize its importance themselves.

3.3 Leo Breiman: Making the Formula "Stand Firm Mathematically"

1961. The University of California, Berkeley.

The Fourth Berkeley Symposium on Mathematical Statistics and Probability was in session. Held once every four years, this symposium was the Olympics of the mathematical statistics world — all the top probabilists and statisticians in the country would come.

At the conference, a statistician named Leo Breiman gave a talk:

Optimal Gambling Systems for Favorable Games

The talk was later collected in the conference proceedings, pages 65–78.

Who was Breiman? The name may be unfamiliar to finance people today, but machine learning people today — everyone who has ever written Python scikit-learn code — surely know him. He later invented two techniques that were pivotal in the history of machine learning:

• CART (Classification and Regression Trees, 1984) • Random Forest (2001)

If you have done machine learning in any form, you have very likely used something of his.

But in 1961, Breiman had not yet done these things. Back then he was just a 33-year-old rising star in statistics, interested in martingale theory and the strong law of large numbers.

He read Kelly's paper. He felt Kelly's derivation was not rigorous enough — Kelly was a physicist, and physicists doing mathematics sometimes think "close enough is good enough." Kelly had proven that "the Kelly strategy maximizes the long-run growth rate," but he had not proven that "the Kelly strategy dominates any other strategy" — that is, he had not proven that "in the long run, any other strategy will lose to the Kelly strategy."

Breiman decided to supply this proof.

3.4 Breiman's Two Theorems

Breiman spent a year using rigorous tools from probability theory to prove two properties of the Kelly strategy. These two properties later came to be called Breiman Property 1 and Property 2.

Property 1 (shortest hitting time):

Under any fixed "wealth target V" (say, "I want to grow my wealth to one million"), the expected time required using the Kelly strategy is the shortest.

This property says: if your goal is "to reach some specific wealth level as fast as possible," the Kelly strategy is the fastest. Any other strategy will make you wait longer.

Property 2 (long-run dominance):

For any essentially different strategy S', the wealth using the Kelly strategy S* satisfies, almost surely:

$$\lim_{t \to \infty} \frac{V_{S^*}}{V_{S'}} = \infty$$

That is to say, in the long run the wealth of the Kelly strategy almost surely dominates any other strategy — the ratio will tend to infinity.

This property sounds very academic, but its meaning is deeply striking:

No matter what other strategy you use, as long as it is not the Kelly strategy, in the long run the ratio of the Kelly strategy's wealth to your strategy's wealth will grow infinitely large.

Not a little more. Infinitely more.

This is the mathematical bedrock of every "long-run optimal" argument for the Kelly formula. Breiman proved it in a 13-page paper.

3.5 An Inconspicuous but Crucial Detail

Breiman's Property 2 contains a phrase — "essentially different." This phrase is extremely important.

It means: if the strategy you use differs from the Kelly strategy only in magnitude (say, you bet 2× Kelly, or half Kelly), then strictly speaking they are not "essentially different" — they both belong to "perturbations of the log-optimal."

But if the strategy you use differs from the Kelly strategy in structure — say, you use "full Kelly but add a 0.5% offset each time," or you use "maximize the arithmetic mean" — then these two strategies are "essentially different."

Property 2 says: all essentially different strategies will, in the long run, be infinitely dominated by the Kelly strategy.

The word "infinitely" in this conclusion is very important. Because it does not require the long-run rate of return to be high — as long as time is long enough, the Kelly strategy wins by an infinite amount.

In the decades that followed, those who opposed the Kelly formula (including Samuelson, whom we will discuss in Chapter 11) would try again and again to refute this Property 2. Their rebuttals were basically all one and the same move: "Winning more in the long run does not mean a rational investor should choose it." — This is a question of preference, not a question of mathematics.

But the mathematical fact is rigorous: in the long run the Kelly strategy almost surely dominates every other strategy. Breiman locked this conclusion down with an airtight proof.

3.6 After Breiman: The Kelly Formula "Stands Firm"

After Breiman's 1961 paper, the Kelly formula formally gained mathematical citizenship.

It was no longer "the clever little trick of an oddball at Bell Labs," but a theorem of probability theory that had been rigorously proven. Later mathematicians — for example, the work of Mark Finkelstein and Robert Whitley in 1981, and the machine learning applications of David Helmbold and Robert Schapire in the 1990s — all cited Breiman 1961 directly as their foundation.

This formula now stood on three legs:

1. Kelly 1956: gave a financial interpretation using information theory (channel capacity = wealth growth rate).
2. Latané 1959: restated it in the language of financial economics (geometric mean maximization).
3. Breiman 1961: rigorously proved long-run dominance using probability theory.

Three legs holding up one table — the Kelly formula was now a table that could actually hold something.

3.7 Where the "Mathematical Beauty" of the Kelly Formula Lies

I want to pause here and talk about the mathematical beauty of this formula.

Many formulas look beautiful but are a nuisance to use — you have to remember a pile of assumptions, look up a pile of parameters, consider a pile of boundary conditions.

The Kelly formula is not like that. Its beauty lies in the extreme simplicity of both its inputs and its output:

• Input: the win probability p (one number), the odds b (one number) • Output: the betting fraction f (one number)

Yet what it tells you is extremely deep:

• It tells you how much you should bet • It tells you why you should not bet more (betting more lowers the long-run growth rate) • It tells you why you should not bet less (betting less wastes your informational edge) • It tells you how to handle estimation uncertainty (bias in parameter estimates is mercilessly amplified by geometric compounding — this was added by later researchers)

On a deeper level: it tells you a core philosophy about uncertainty —

How much information you have determines how much you can win.

This is the most profound sentence of the Kelly formula. It places information and wealth on the two sides of the same equals sign.

If you have no information — your win probability is only a hair above 50% — then the Kelly formula tells you to bet very small. If you have abundant information — your win probability is significantly above 50% — then the Kelly formula tells you to bet big. If you have no information at all — the win probability is exactly 50% — then the Kelly formula tells you not to bet.

This is a very honest formula. It does not allow you to deceive yourself.

In later practice, all the "traps" of the Kelly formula are, in essence, because people deceived themselves — overestimating their own win probability, underestimating uncertainty. But as long as the win probability you input is honest, the betting fraction the Kelly formula outputs is mathematically optimal.

3.8 Latané, Breiman, Kelly: Three "Quiet Founders"

I want to pause once more here and compare the three founders.

John Kelly: a PhD in physics, a former pilot, dead of a stroke at 41. He never saw the impact of his formula. The most famous work he left behind at Bell Labs was, ironically, "making a computer sing."

Henry Latané: a professor of financial economics at the University of North Carolina, died in 1996. He may have gone his whole life without realizing how important the 1959 paper he wrote was.

Leo Breiman: a statistician, a professor at Berkeley, died in 2005. He later became famous in the machine learning world for inventing CART and Random Forest. In his obituary, the Kelly formula was mentioned in only a single line.

Three men, three fields, three points in time: physics 1956, economics 1959, statistics 1961. All three had no idea what the others were doing. All three did not live to see their work change the world.

But — taken together, the three of them gave the Kelly formula a heart that could beat.

The physicist gave it intuition (channel capacity = wealth growth). The economist gave it a financial identity (geometric mean maximization). The statistician gave it a rigorous soul (long-run dominance).

After this, the formula was about to leave academia and step into its true destiny —

The casino. Wall Street. The racetrack.

See you in the next chapter.

Chapter Summary: The "Tripod" of the Kelly Formula Founder Year Contribution Key Concept John L. Kelly Jr. 1956 Proposed the formula Channel capacity = wealth growth rate Henry Latané 1959 Introduced it to financial economics Geometric mean maximization Leo Breiman 1961 Rigorous mathematical proof Long-run dominance, shortest hitting time

Part One, "Birth," ends here. In Part Two, we will see this formula step out of academia and into the real casinos and Wall Street — and the first person to bring it in was that young mathematics lecturer who, in 1960, took a sheaf of paper from Shannon's drawer: Edward Thorp.

Part Two · Toward the Real World

Chapter 4 · A Mathematics Professor's Night in Las Vegas

4.1 Ten Thousand Dollars, Three Casinos, Seventy-Two Hours

The early morning of a weekend in 1961. Reno, Nevada.

A freshly rented Chevrolet was parked in the lot outside the casino. Three men sat inside:

• Edward Thorp: twenty-nine years old, a mathematics lecturer at MIT, fair-haired, wearing black-framed glasses, looking like a graduate student who didn't much like to talk. • Emanuel "Manny" Kimmel: a New Yorker in his early sixties, immaculately suited, never without a cigar. His "legitimate business" was running the Kinney parking-lot chain—the very company from which the Warner Bros. you know today would evolve. But his "side business"… let's leave that unsaid for now. • Eddy Hand: Kimmel's "partner," and one of the financial backers.

In the trunk of the car sat ten thousand dollars in cash—mostly hundred-dollar bills, bundled into a hundred stacks.

Converted to 2026 purchasing power, those ten thousand dollars would be worth about a hundred and ten thousand dollars today.

The three men walked into the casino. Thorp took a seat at the blackjack table. Kimmel and Hand stood apart on either side behind him—the posture of gamblers, and also the posture of bodyguards.

Seventy-two hours later, they walked out of the casino. Net winnings: eleven thousand dollars—which is to say, they had more than doubled their bankroll.

At 2026 purchasing power, that amounts to a net gain of a hundred and twenty thousand dollars in seventy-two hours.

But the more important thing is not this number. The more important thing is this—

This was the first time in human history that anyone had beaten a commercial casino using mathematics.

4.2 Who Was Thorp

To make the weight of this event clear, we must first make clear who Thorp was—and how he came to blackjack.

Edward Thorp was born in Chicago in 1932. His father was a World War I veteran who had worked as a warehouse manager. The family was very poor. Thorp's only toy as a child was the library—by the age of sixteen he had read through all the mathematics and physics books in the town library.

At twenty-one he earned his bachelor's degree from UCLA (the University of California, Los Angeles), and at twenty-four he received his doctorate in mathematics from UCLA, specializing in functional analysis.

In 1959, he came to MIT as a mathematics lecturer.

That year he was twenty-six.

Many people think of Thorp as a gambler. In truth he was no gambler at all—he did not enjoy the thrill that gambling brought, and he did not even like going to casinos. He took up blackjack purely because he thought the claim that "the casino cannot be beaten" was wrong—anything declared "impossible" is, to a mathematician, a challenge.

In his final year of doctoral study at UCLA, he heard of a 1956 paper titled "Optimal Strategy for Blackjack," written by Roger Baldwin and three other researchers. Using hand-cranked calculators (in an era before IBM machines were common), this paper had worked out the "basic strategy" of blackjack—telling you when to hit, stand, double down, or split on every hand.

Baldwin's work proved that: playing blackjack with basic strategy, the house edge was only about 0.5%—very close to a draw, but still a slight advantage for the house.

Thorp's first reaction on seeing this paper was: "These fellows missed the key point—the composition of the remaining deck."

The key to blackjack is this:

• If the remaining cards contain a high proportion of high cards (10, J, Q, K, A), the player has the advantage. Because the dealer is more likely to bust, and the player is more likely to get a 21. • If the remaining cards contain a high proportion of low cards (2–6), the dealer has the advantage.

Every time a card is dealt, the composition of the remaining deck changes once. So in theory, the player can dynamically adjust strategy and bets according to the "remaining deck."

This is what later became the famous "card counting."

4.3 How Card Counting Was Invented

Thorp knew the idea, but to actually calculate it out—to produce a system usable in real play—required computing power far beyond a hand-cranked calculator.

Fortunately, in 1959 MIT had an IBM 704—one of the most advanced scientific computers in the world at the time. Thorp did not know how to program. But by then the IBM 704 already had a new language available, called FORTRAN (short for "Formula Translation," invented by IBM in 1957). Thorp taught himself FORTRAN over a few weeks, wrote a few hundred lines of code, and had the IBM 704 simulate several million rounds of blackjack, calculating the player's optimal strategy and true win probability under every kind of "remaining-card composition."

After running for several days and nights, the result came out:

When the proportion of high cards was high enough, the player's edge could exceed 1%.

One percent does not sound like much. But this is the extreme limit of the maximum edge a casino allows a player—any edge above 1% means the casino is bound to lose. If you can bet big when the player has the advantage and bet small when the dealer has the advantage, you are sure to win in the long run.

But how to know the composition of the remaining deck? Count 52 cards? Count six decks, 312 cards in all? The human brain cannot do it.

Thorp came up with a clever simplification—the "Five-Count System":

Just count how many 5s are left. If all the 5s are gone, the player clearly has the advantage.

Later he refined it into a more general "Ten-Count"—counting the remaining proportion of 10 / J / Q / K.

Later still (in 1963), working with the engineer Harvey Dubner, he invented what is today the most famous counting method, "Hi-Lo":

• 2–6 count +1 (low cards) • 7–9 count 0 (neutral cards) • 10–A count −1 (high cards)

Your "running count" is the sum of the counts of every card you have seen. The higher the count, the more high cards remain in the deck, and the greater the player's advantage.

Any ordinary person can learn to use this system in real time at the table after a few months of practice—and with enough precision to beat the casino.

4.4 But Counting Tells You "How Much Edge"—How Much Do You Bet?

In 1960 Thorp published that NAS paper Shannon had recommended, "Fortune's Formula: A Winning Strategy for Twenty-One"—and the phrase "Fortune's Formula" in the title came precisely from the Kelly formula.

But this NAS paper had not yet truly made use of the Kelly formula. It spoke of how to obtain an edge, not of how to allocate the bankroll on each hand.

It was precisely into this gap that, in November 1960, Shannon handed him that sheaf of Kelly's papers.

Once Thorp had read and understood it, the chain of logic ran like this:

1. Card counting tells me what my win probability is at a given moment (say 51%, 52%, 53%).
2. The odds on each hand are roughly 1:1 (close to it).

3. Substitute into the Kelly formula f* = (bp − q)/b:

True count TC Player edge Kelly bet fraction f* 0 (neutral) -0.5% (dealer advantage) no bet +1 0% no bet +2 +0.5% about 0.4% +3 +1.0% about 0.8% +4 +1.5% about 1.2% +5 +2.0% about 1.5%

(Here an approximate formula f* ≈ Edge / Variance is used, because the payoffs in blackjack are not simple binary outcomes—there are ties, there is the 1.5× payout on blackjack, there is doubling down, there is splitting.)

The Kelly formula gave Thorp a complete betting system:

• Count ≤ +1: bet the minimum (both protecting the bankroll and disguising the behavior so the casino would not grow suspicious) • The higher the count, the larger the bet • Cap on any single bet: 1% of total capital (far below what the Kelly formula advised—because from the very start Thorp used "half-Kelly" or less)

Put together, this system was called the "Thorp System," a watershed in the history of blackjack gambling.

4.5 Finding the Money

The theory was in hand. The next step was to find ten thousand dollars for a bankroll—by the prices of 1961, that was no small sum.

Thorp was an MIT lecturer, with an annual salary of only a little over nine thousand dollars. He could not come up with the ten thousand himself.

He needed a backer.

The opportunity came. Around Christmas of 1960, after the NAS paper was published, Thorp gave several media interviews. The Washington Post and Time magazine both reported on him—"MIT Mathematician: I Beat Blackjack With Math."

The day after the reports ran, the phone at Thorp's home rang off the hook. Dozens of calls a day—people who wanted to learn counting from him, people who wanted to mail him money to gamble on their behalf, people who wanted to offer him a book deal—and among them a few unusual calls.

One day, at the other end of the line, came the voice of a man, coarse and slow:

"Professor Thorp, I read your paper. I have money. Do you want to give it a try?"

That man was Manny Kimmel.

4.6 The Man Named Kimmel

Manny Kimmel—a name today's financiers may not have heard, but the things he did could fill a book of their own.

He was born in New Jersey in the early 1900s, and amassed his first fortune before the Great Depression—through bootlegging, underground lotteries, and bars. During America's Prohibition (1920–1933), he had close business dealings with the mob (including "Lucky" Luciano's New York crime family).

After Prohibition ended, he went straight—turning his business toward a parking-lot chain (the Kinney Parking System). That company was later acquired in the 1960s and became the Kinney National Company of today. Kinney then acquired a company called Warner Bros. Records, and eventually evolved into the forerunner of today's Warner Bros. Discovery.

But Kimmel never washed away his "gray background." He always lived in a world of blurred boundaries—one foot planted in legitimate business, the other still dusted with the grime of an older era.

In 1961 he was in his sixties, and one of the things he loved was gambling. He had thrown millions of dollars around casinos, won and lost. He knew better than anyone: the casino cannot be beaten.

Until he saw Thorp's paper.

He told Thorp over the phone: "I'll put up ten thousand dollars for the bankroll. We split it four-six—four for me, six for you. You focus on playing the cards; I'll handle… the other things."

Thorp hesitated for a long time. He knew Kimmel's background—he did not like dealing with such people. But he needed the money.

More important still—he genuinely wanted to know whether his theory would hold up inside a casino.

He agreed.

4.7 Those Seventy-Two Hours

On a weekend in 1961, Thorp, Kimmel, and Hand flew to Reno.

The first night, they went to Harold's Club—one of the largest casinos in Reno.

Thorp took a seat at the blackjack table. He looked nothing like a professional gambler—he wore the suit of an MIT professor, wore his glasses, and looked nervous and stiff. For the first few hands he bet the minimum, one dollar.

In his head he was counting the Hi-Lo count.

Slowly, the count climbed. +3, +4, +5. He raised his bet from one dollar to a hundred dollars—a hundredfold jump on a single hand.

Behind him, Kimmel broke out in a cold sweat.

At that table, Thorp won several thousand dollars.

The second night, Cal-Neva Lodge (on the shore of Lake Tahoe, on the border of California and Nevada, where Frank Sinatra and the Kennedy family had both played).

The third night, the Sands.

Over three days and three nights, the three men's eleven-thousand-dollar bankroll turned into twenty-one thousand. A net gain of eleven thousand—which, calculated at 2026 purchasing power, is a hundred and twenty thousand dollars.

4.8 The Casino Reacts

But more dramatic still—the casino began to react.

On the second night at Cal-Neva, the dealer began shuffling frequently—reshuffling after every two hands. This destroyed the premise of a "stable remaining deck" and rendered card counting useless.

Thorp knew this was the casino "counter-attacking"—they had noticed an odd man in a suit winning money.

On the third night at the Sands, a cocktail waitress brought Thorp a glass of whiskey—Thorp had not ordered it. He took a sip, and his head began to spin.

He suspected—it may well have been true—that the whiskey had been drugged.

Thorp made an excuse to go to the restroom and poured the drink away. He switched tables and kept playing.

But the casino's expression had already changed.

On the morning of the fourth day, Kimmel said: "Let's go. These people will do worse things."

The three men drove out of Reno, the eleven thousand dollars tucked in the inner pocket of Kimmel's suit. On the way back to Nebraska, they stopped at a gas station. Kimmel finished a cigar and said to Thorp:

"Professor, you did it. It's real. I've gambled my whole life, and I've never seen anything like this."

Thorp later wrote in his autobiography A Man for All Markets (2017): in that moment, for the first time, he fully believed in the Kelly formula—not because the formula was beautiful, but because it had truly won money inside a casino.

4.9 Beat the Dealer Is Published

In 1962, Thorp wrote this system—card counting plus Kelly betting—into a book called Beat the Dealer: A Winning Strategy for the Game of Twenty-One.

Published by Random House.

It sold five thousand copies in the first week. Fifty thousand in the first month. Two hundred thousand in the first year.

It was the first bestseller in the history of gambling. It set off three chain reactions:

First: tens of thousands of ordinary people bought the book and went off to Las Vegas to put card counting into practice. Most of them executed it very badly—counting is physically demanding, and after a few hours one's concentration collapses. But a few succeeded—such as the later MIT blackjack team (the film 21 that you watch today is adapted from their story).

Second: the casinos were forced to reform.

• Beginning in 1965, most casinos changed from single-deck to multi-deck games (4 decks, 6 decks, 8 decks), making counting harder. • Beginning in the 1970s, "shuffling machines" became common, so that every hand started from a fresh deck and counting was rendered useless. • Casinos established "blacklists"—photographs of counters were circulated across the whole country, and players who were caught were banned for life. • Some casinos used a "reshuffle at any time" rule—the dealer could shuffle at any moment, breaking up the count.

Third: from then on, the Kelly formula had a name. In the 1966 revised edition, Thorp formally named it "the Kelly Criterion"—and this term has been fixed ever since.

4.10 Why Thorp Left the Casino

After playing blackjack for a few years, Thorp voluntarily withdrew.

Not because he could not afford to lose. Nor because he had been barred by the casinos (he could still get in afterward—as long as he changed his face, wore a false mustache, and booked rooms under a different name).

It was because he realized one thing:

There is a ceiling on the money blackjack can win.

A casino's cap on a single hand was usually five hundred dollars, at most a thousand. One person can play a hundred rounds in a night. Even with perfect counting, the expected profit per hour is only fifty to a hundred dollars. Over a year—a few tens of thousands of dollars—and he could earn that much in a year as an MIT professor too.

But the financial markets have no ceiling.

If the same "Kelly spirit"—finding an edge, betting sensibly, compounding over the long run—could be applied to the stock market, then what he earned in a year might be hundreds of thousands, even millions.

In 1962, Thorp left MIT and went to New Mexico State University as an associate professor. In 1965 he went to the University of California, Irvine (UC Irvine) as a full professor. From then on, he taught mathematics by day and studied finance by night.

What was he studying? Options.

That was 1965—no one had yet published a formula for pricing options. The Black-Scholes formula would not be published until 1973.

But Thorp—and his colleague in New Mexico, Sheen Kassouf—had, by 1967, using pure hedging reasoning, independently guessed at an option-pricing formula almost exactly equivalent to Black-Scholes.

They titled the book Beat the Market.

In 1969, Thorp partnered with a Wall Street stockbroker, Jay Regan, and opened a small hedge fund.

That fund was first called Convertible Hedge Associates. In 1974 it was renamed—

Princeton-Newport Partners.

The story of this fund over the next nineteen years is the first formal appearance of the Kelly formula on Wall Street.

See you in the next chapter.

Chapter Summary: What Thorp Proved at Blackjack

First: theory can hold up in the real world. The Kelly formula is not a blackboard toy—it really can win money in Las Vegas.

Second: the correct use of the Kelly formula is conservative. Even though Thorp could calculate that Kelly advised betting 1.5%, in real play he used a figure far smaller than that—no single bet exceeding 1% of total capital. This was the seed of the later idea of "fractional Kelly."

Third: commercial casinos are not invincible. But the price of beating the casino is being counter-attacked by it—shuffling, blacklists, drugs in the drink. This story tells us: any edge is temporary; the market (or the casino) will react. This rule still holds seventy years later among quantitative hedge funds.

Fourth: the profit ceiling of blackjack is too low. Thorp saw this clearly, and turned his gaze to the stock market—a world with a far higher ceiling, but where the edge is far harder to find.

In the next chapter, what we will see is his first move in that world.

Chapter 5 · Wall Street's Quietest Nineteen Years

5.1 A Miracle No One Cared About

If you open up the history of American hedge funds and search for "the most stable hedge fund of all time," the vast majority of answers will give you two names:

• Renaissance Technologies' Medallion Fund (founded by James Simons, 1988–present) • Princeton-Newport Partners (founded by Edward Thorp, 1969–1988)

The Medallion Fund is the more famous of the two — over more than three decades it has earned an annualized net return exceeding 40%, earning it the reputation of "the greatest hedge fund of all time."

But Princeton-Newport Partners (hereafter PNP) — across the entire 19 years of its existence — accomplished something even "quieter," and even harder to replicate, than Medallion:

Not a single losing quarter.

Not merely not a single losing year — not a single losing quarter. 19 years × 4 quarters = 76 quarters, every one of them a positive return.

By the public figures:

• Gross annualized return of about 19.1% • Net annualized return of about 15.1% after deducting all fees • Over the same period, the S&P 500's net annualized return was about 10% • Over the same period, its volatility was only 1/3 that of the S&P 500

But even more astonishing was its performance during times of crisis:

• The 1973–74 bear market: the S&P 500 fell 41% over two years. PNP still rose in both of those years — up 6.5% and 9% respectively. • Black Monday, October 1987: the S&P dropped 22% in a single day, the worst day since the Great Depression. PNP was essentially flat that month.

19 years, 76 quarters, every one a win.

This is not the kind of "luck" you commonly find in markets. This was the first formal appearance of the Kelly formula on Wall Street.

5.2 Thorp + Regan: An Odd Pairing

The story of PNP has to begin with the meeting of two men.

1969. California.

Edward Thorp — you already met him in Chapter 4 — was 37 at this point, a full professor in the mathematics department at UC Irvine. After the publication of "Beat the Market" in 1967, his name had spread through Wall Street circles. Some people who "knew a good thing when they saw it" realized that the hedge arbitrage methods in this book — far predating the Black-Scholes formula — were already capable of turning a profit in real trading.

Jay Regan — a 27-year-old Philadelphia securities broker, an undergraduate of Dartmouth College who had spent a few years working as a stockbroker on Wall Street. He read "Beat the Market" and flew to California to find Thorp.

He said to Thorp: "I've studied your theory. It can make money. But you don't understand Wall Street — you don't know how to find clients, how to haggle with brokers, how to handle taxes, how to deal with regulators. I know all of that. Let's be partners."

Thorp hesitated. He didn't like Wall Street, and he didn't like leaving California.

Regan said: "You don't have to move. You do the research, the trading, and the risk control in California. I'll handle clients and execution in Princeton, New Jersey. We'll coordinate by telephone and telegraph."

— This was 1969. There was no internet, no email; even the fax machine was not yet widespread. Two men coordinated a hedge fund by telephone and teleprinter, one in California, the other in New Jersey.

But they pulled it off. In November 1969, Convertible Hedge Associates was established. Thorp made the investment decisions; Regan handled clients and execution.

In 1974 it was renamed Princeton-Newport Partners — "Princeton" referring to Regan's office in Princeton, New Jersey, and "Newport" referring to Thorp's office in Newport Beach, California.

Two offices in two locations. A collaboration across the Western and Eastern time zones.

The first real-world battle of the Kelly formula on Wall Street began just like this.

5.3 PNP's Core Strategy: Convertible Bond Arbitrage

What PNP did sounds more complex than blackjack — but in essence it was the financial version of blackjack:

Find hedge trades with positive expected value, then allocate capital using the Kelly formula.

PNP's most profitable strategy was called "convertible bond arbitrage." Put simply:

• A convertible bond is a special kind of bond. It pays you interest (like an ordinary bond), but you can choose to convert it into stock (like an option). • Therefore the price of a convertible bond = the value of an ordinary bond + the value of an option. • If the market prices the convertible below the sum of these two parts — then you can go long the convertible bond and short the corresponding stock, locking in a risk-free profit.

The key to this strategy is calculating that "option value" correctly. As early as 1967 Thorp had independently worked out a pricing formula similar to Black-Scholes — so he knew which convertible bonds were undervalued.

For each trade, the expected profit on a single position might be only 1–2%, but the risk was also very small — because the stock's direction had already been hedged away. This is precisely the scenario the Kelly formula loves: a small edge, low risk, and the ability to repeat at high frequency.

PNP made dozens, even hundreds, of such small trades every day. Each one, viewed on its own, was nothing much, but added together — a long-term annualized return of 19%.

5.4 PNP Used "Half Kelly"

There was one iron rule in PNP's operations:

No single position was to exceed 2% of total capital.

Why 2%? Because Thorp had calculated: • Even if one of his trades was 100% wrong about direction (the worst case), the most he could lose was 2% • Out of 100 trades, even if 10 were completely wrong, the most he could lose was 20% — far below the fatal level • And if, out of 100 trades, 90 were small wins — the law of large numbers would guarantee the portfolio's profitability

At a deeper level, the 2% cap was a discount on the Kelly formula.

In theory, many of PNP's convertible bond arbitrage trades would, under full Kelly, warrant betting 5%–10% of capital. But Thorp deliberately cut that figure down to 1/3–1/5 of the original — that is, he used "1/4 to 1/2 Kelly."

Why cut it this way? There were three reasons:

First: parameter estimation error. Any estimate of "option value" carries a ±10% margin of error. Fed into the Kelly formula, this error turns "full Kelly" into what is actually "over-Kelly" — which loses money over the long run. Cutting Kelly in half completely absorbs a 10–20% estimation error.

Second: liquidity risk. The convertible bond market is less liquid than the stock market. Once a crisis hits, if you want to exit, you may not be able to sell at all. Cutting the position is equivalent to leaving yourself room to "ride it out."

Third: psychological risk. Even if your mathematics is correct, after several consecutive losses your psyche will break down — you'll cut your losses and get out, and miss the rebound. Cutting the position is equivalent to leaving room for the psyche.

Later these three reasons — estimation error, liquidity, and psychology — became the shared understanding of all professional Kelly users. When we look at Bill Benter at the racetrack in Chapter 9, and at Kelly systems in the age of AI in Chapter 13, we will see these three again and again.

5.5 A Forgotten Detail: Thorp vs. Black-Scholes

Writing this far, I want to insert an aside.

In 1973, Fischer Black and Myron Scholes published in the "Journal of Political Economy" the famous paper "The Pricing of Options and Corporate Liabilities" — this is the source of the Black-Scholes option pricing formula. It won Scholes the Nobel Prize in Economics in 1997 (Black had already died in 1995, and by the rules the prize cannot be awarded posthumously).

This formula has been hailed as "the most important formula in 20th-century finance," the cornerstone of derivatives pricing.

But —

As early as 1967, Thorp had independently guessed at an almost entirely equivalent formula. In Chapter 6 of "Beat the Market," using a kind of pure geometric reasoning called "delta hedging," he wrote out an approximate formula for the value of an option. The structure of this formula was almost identical to Black-Scholes's 1973 formula — differing only in the handling of some boundary conditions.

Why didn't Thorp publish his "complete version"? Two reasons:

First: he didn't want to make it public. He was using this formula in real trading to make money (after PNP was founded in 1969). If he published it, there would be no more money to make.

Second: he was not a mathematical economist. What he did was practical derivation — good enough to use, without pursuing a rigorous proof in continuous-time stochastic processes. Black-Scholes used full Itô calculus and partial differential equations — an academically more rigorous proof.

So the Nobel Prize went to Scholes and Merton. Thorp, with a clear conscience, went on making money in California.

Later, in his 2017 autobiography, Thorp said modestly: "Black-Scholes's 1973 formula was far more rigorous and far more beautiful than my 1967 version. I concede completely. But back in the day, the man who first used it to make money was me."

This is one of the deep spirits of the Kelly formula: making money does not require a Nobel Prize. What it requires is honestly calculating your edge, allocating your position sensibly, and executing repeatedly over the long run.

5.6 Black Monday, 1987: The Kelly Formula Stands the Test

October 19, 1987. Monday.

The Dow Jones Index fell 22.6% in a single day. This was the worst day for the American stock market since the Great Depression of 1929. Everyone was bleeding.

Many hedge funds — especially those using leverage — saw 30%, 40%, even 50% of their capital evaporate on that single day. Some of them never recovered.

How did PNP do that day?

That day PNP had a small loss — but for the month as a whole it was essentially flat.

Why?

First: PNP's strategy was market-neutral — long the undervalued convertible bonds, short the corresponding stocks. The market fell 22%, and both sides fell — but PNP's "spread position" barely moved.

Second: PNP's positions were diversified — hundreds or thousands of small hedge positions. No single position exceeded 2%. Even if some positions ran into trouble, the portfolio as a whole would not suffer a major loss. Third: PNP's leverage was measured — following the discounted version of the Kelly formula. Its "gross book exposure" might have been 5–8 times capital (a typical hedge fund level), but its net exposure — that is, the part actually facing market volatility — was very low.

On October 19, LTCM had not yet been founded (that would not come until 1994) — later we will see that LTCM used 25–30 times leverage and collapsed completely in 1998.

PNP used the discounted version of the Kelly formula. LTCM did not.

The difference between the two is the fundamental reason PNP survived until 1988 while LTCM died within three months.

5.7 The End of 19 Years: A Legend Broken Open by the Police

PNP did not die from the market.

It died from a judicial incident.

December 1987 — two months after Black Monday — the FBI raided PNP's office in Princeton.

The origin of the affair was this: the office on Regan's side had extensive trading dealings with Drexel Burnham Lambert (Drexel — the most controversial investment bank on Wall Street at the time, the firm where Milken worked). While the FBI was investigating Drexel's "junk bond" insider trading case, following the trail led them to suspect that PNP was involved in a violation called "stock parking" — helping Drexel's clients avoid taxes through sham trades.

Thorp — on the California side — knew nothing of these matters at all. He taught mathematics by day and did convertible bond arbitrage by night. What was happening on the Princeton side he heard only in fragments over the telephone.

After a year of investigation and judicial proceedings, PNP was charged with a number of violations (most of them administrative, involving neither violence nor major fraud). At the end of 1988, Thorp decided to shut PNP down — he was unwilling to bear the cost of these administrative lawsuits.

In December 1988, PNP was formally wound up.

19 years. 76 consecutive positive quarters. Never once a losing quarter.

And in the end it was broken open by the police.

5.8 Thorp's Second Spring: Ridgeline Partners

But the story does not end there. Thorp was 56 in 1988. He rested for 5 years — returning to California, writing papers, doing some consulting work, living a retired life for a while.

In 1994, he reopened a hedge fund called Ridgeline Partners. This time there was no Regan — Thorp led it alone.

Ridgeline's strategy was primarily statistical arbitrage — a method Thorp had already developed in PNP's later years. It was the same class of strategy as Renaissance's Medallion, only smaller in scale and more discreet.

Ridgeline's net annualized return from 1994 to 2002 was about 21%. Its Sharpe ratio was 1.88 — a world-class level.

In 2002, Thorp voluntarily shut Ridgeline down — he said, "the market has gotten crowded, the edge is disappearing." He did not wait for the edge to disappear before closing — he closed while the edge was still there.

This too is a kind of Kelly spirit — when you know your edge is no longer reliable, stop immediately.

Over his whole life, Thorp managed two hedge funds in total, 28 years combined, with never a losing year.

By today's standards, his personal net worth is estimated at around 800 million dollars. But his influence far exceeds that figure — he directly or indirectly trained the next generation of quantitative investors, including Renaissance's early employees, AQR founder Cliff Asness, and part of Citadel's team.

More importantly — he gave the Kelly formula evidence on Wall Street. Before him, the Kelly formula was merely an academic theory. After him, the Kelly formula became a standard tool of hedge fund risk control.

5.9 The Counsel of Thorp's Later Years

In 2017, Thorp, at 85, published his autobiography "A Man for All Markets." In this book he wrote a passage of counsel to the ordinary investor.

I translate it here word for word:

"I'm often asked: 'Should I use the Kelly formula?'

My answer is: yes, but use 'half Kelly' or lower.

Because you do not know your own win rate. You think your win rate is 55% — it may actually be 50%. You think it's 60% — it may actually be 53%. Everyone overestimates their own win rate, including me.

Half Kelly keeps you safe even under estimation bias. Full Kelly is optimal only for the perfect oracle — but the perfect oracle does not exist. In all my life of investing, I have never used full Kelly on a single major trade. I have lived to 85, and my account has lived to 85 too. The two are one and the same thing."

Chapter Summary: What PNP Left Behind

First: the Kelly formula can be used in real trading on Wall Street — as long as what you do is hedge trading (market-neutral, repeatable, with a quantifiable edge).

Second: in real trading, never use full Kelly — use 1/4 to 1/2 Kelly. This rule is not conservatism; it is honesty about your own estimating ability.

Third: never let a single position exceed 2% of capital. This is an iron rule blackjack brought to Wall Street — Bill Gross applied this rule at a scale of 2 trillion dollars in the bond market (next chapter).

Fourth: when the edge disappears, stop immediately. Ridgeline's voluntary shutdown in 2002 was exactly this principle. Stopping before the edge disappears is better than stopping only after the edge disappears.

In the next chapter, we will look at a man completely unlike Thorp — he never read the NAS papers, never taught at MIT, never chatted with Shannon. He was a psychology undergraduate at Duke University who, on the eve of graduating in 1966, having read "Beat the Dealer," went to Las Vegas and played blackjack for 4 months, turning 200 dollars into 10,000.

Then he used that experience to manage 2 trillion dollars in bonds.

His name is Bill Gross.

Chapter 6 · From Two Hundred Dollars to the Bond King

6.1 A Duke Psychology Student's Car Crash

Spring, 1966. Durham, North Carolina. The campus of Duke University.

A twenty-two-year-old senior was driving a beat-up Ford back to his dormitory along some highway just off campus.

His name was William Hunt Gross — his friends called him Bill. He was a psychology major, about to graduate, with no particular plan for his life. His family was ordinary; he had no rich father, no network of connections reaching out a helping hand.

That night, his car ran off the road and flipped. Bill was thrown clear of the vehicle. His skull was fractured, and he was bleeding inside his head. He lay in the hospital for months — for the first six weeks he could not even speak or read. The doctors told his parents that his surviving at all was a miracle, and that his recovering to something close to a normal person would be a still greater miracle.

— And the miracle really did happen. Bill slowly recovered.

But one thing was changed forever:

He no longer believed in "the life lived step by step."

He had nearly died. Death had looked him over at close range. From that moment on, he wanted to do something that would excite him — and "following the path his teachers pointed to, on to graduate school and a steady job in psychological counseling" was not one of those things.

The first thing he did that excited him was this book:

Beat the Dealer by Edward O. Thorp.

He read it in his hospital bed.

6.2 The Two-Hundred-Dollar Las Vegas Experiment

Summer, 1966. Bill graduated from college.

He made a decision: with two hundred dollars as his bankroll, he would go to Las Vegas and play blackjack, to see whether Thorp's system was really true.

Two hundred dollars, in 1966 purchasing power, was the equivalent of about eighteen hundred dollars today — a sum neither large nor small. To a young man just out of school, with no income, it was very nearly his entire savings.

He didn't tell his parents. He bought a cheap plane ticket to Las Vegas and checked into a run-down hotel downtown that cost eight dollars a night.

Over the next four months, he did something almost impossible to imagine —

He played blackjack sixteen hours a day.

Eight in the morning until noon. Two in the afternoon until eight at night. Ten at night until two in the morning. Sixteen hours a day, four months without a break.

What he used was exactly what Thorp's book had taught — "card counting plus fractional-Kelly betting" — the method handed straight down from the Kelly-Thorp lineage. Four months later, his two hundred dollars had become ten thousand — a fiftyfold increase.

At today's prices, ten thousand dollars is equivalent to about ninety thousand. A twenty-two-year-old unemployed youth had made ninety thousand dollars in four months.

This event changed the whole of his life.

Not because of the ten thousand dollars — that little sum was nothing against the life that lay ahead of him. It was because he had verified for himself, in his own flesh, a principle:

If you have a real edge (however small), and you bet strictly according to a sensible proportion, over the long run it will accumulate into an enormous fortune.

This principle — together with what the crash had taught him, that "life is short, and you must do what excites you" — became the two North Stars of his entire later life.

6.3 From Blackjack to the Bond Trading Desk

After winning at the tables, Bill served a brief stint in the Navy (during the Vietnam War), and then flew to the West Coast to earn an MBA at UCLA's Anderson School of Management.

Why UCLA? Because he paid part of his tuition with the money he had won gambling — California's state universities were relatively cheap — and because he did not want to be too close to New York and Wall Street; he wanted to find his own way into this business of investing.

Graduating with his MBA in 1971, he joined an insurance company called Pacific Mutual Life Insurance, working in its bond investment department.

In those days bond investing was still a lifeless industry. The job of a bond manager was to "buy and hold" — to buy some high-rated corporate bonds and government bonds, collect the interest each year, get the principal back at maturity, and never actively trade. The whole industry was old men managing money — average age fifty and up.

On his first day there, Bill's boss told him:

"Son, bonds are just like this — you buy, you hold, you collect the interest. There's nothing fancy to play with here."

Bill disagreed.

His head was full of blackjack — where every hand demanded thought, every hand demanded adjustment, every hand demanded a dynamic decision based on the "remaining deck." What if the bond market could also be traded dynamically, like blackjack — adjusting the position dynamically according to changes in interest rates, credit spreads, and differences in liquidity — what then? He began to trade bonds actively — buying and selling when no one else was trading. He hunted for bonds the market had mispriced — just as, in blackjack, he had hunted for the moments when "many high cards remained."

What position rule did he use? The same rule blackjack had taught him:

No single position exceeds 2% of total capital.

This rule would later become the iron law of his entire financial life.

6.4 The Birth of PIMCO

In 1971 — the same year Bill joined Pacific Mutual — Pacific Investment Management Company was founded. This was the subsidiary Pacific Mutual set up specifically to manage bonds, PIMCO for short.

Bill was among PIMCO's first employees, and its de facto head of bond investment.

At that time PIMCO managed very little money — just a few million dollars when it was founded, mostly Pacific Mutual's own internal funds.

But Bill's strategy was different. He managed bonds actively, using "the discipline of blackjack" for risk control:

• No single issuer exceeds 2% of total assets (this rule was later extended to "no single credit entity exceeds 2%") • Total leverage no more than 1.5 times (far below the industry's common 5-to-10 times) • Every trade first asks "what happens if this is completely wrong" (this is the "Gambler's Ruin" concept from blackjack — you cannot allow any single trade to bankrupt you)

He traded actively — constantly adjusting the portfolio according to interest-rate expectations, credit spreads, the shape of the yield curve, and the distribution of bond maturities.

Ten years later, PIMCO's assets under management had grown from a few million to a few billion dollars. Twenty years later (1991), PIMCO managed sixty billion dollars. Thirty years later (2001), PIMCO managed two hundred and fifty billion dollars. Thirty-five years later (2006), PIMCO managed six hundred and sixty-eight billion dollars. Forty years later (2010), PIMCO managed more than one trillion dollars. Forty-four years later (2014) — the year Bill was about to leave — PIMCO managed two trillion dollars — the largest bond fund management company in the world.

Bill was repeatedly named by the Financial Times, Barron's, and The Wall Street Journal as "the greatest bond investor of all time."

He earned himself a nickname — "the Bond King."

6.5 The Miracle of the PIMCO Total Return Fund

PIMCO's most famous product was called the PIMCO Total Return Fund — PTTRX for short.

From its founding in 1987 until Bill's departure in 2014, this fund long outperformed the Barclays Aggregate Bond Index by about 150-200 basis points (1.5%-2.0% a year).

Doesn't sound like much?

But it managed two hundred billion dollars. 1.5% × 200 billion = 3 billion dollars — this is the wealth the PIMCO Total Return Fund created for its clients over and above the index each year through active management. Compounded over twenty-seven years, more than eighty billion dollars.

This fund became one of the bond funds most commonly held by American households — from middle-class retirement accounts to university endowments, from local-government pensions to corporate cash management, nearly every American directly or indirectly held PIMCO's bonds.

And the foundation beneath all these miracles — as Bill said countless times in interviews — was the rule blackjack had taught him:

Never let any single bet exceed 2% of capital.

6.6 Why Bill's "2% Rule" Is the Essence of the Kelly Formula

Many people — including many bond practitioners — think Bill's "2% rule" is empiricism. Bill himself often said as much: "I learned it from blackjack."

But the mathematical root of this rule is precisely the Kelly formula.

Let us take a look:

In bond investing, the "default risk" of a single issuer is usually far lower than that of a stock. The annual default rate on investment-grade corporate bonds is about 0.1-0.5%. That is to say, the probability of a loss on any single investment is very low, but once it happens, the size of the loss can be very large (the recovery rate after default is about 40-60%).

Plug into a simplified Kelly formula:

• Win rate p = 99.5% (no default) • Loss rate q = 0.5% • If you win, you earn 5% (the interest-rate premium) • If you lose, you lose 50% (you recover 50% of principal after default)

By the continuous version of the Kelly formula — the theoretically optimal position ratio would come out very high (the mathematically optimal solution, under the assumption of "perfectly accurate parameters," could exceed 50% or even more). But this is exactly the crux of the problem — this assumption of "perfect accuracy" almost never holds in bond investing.

But —

Bill knew two things:

First: his estimate of any single issuer's default probability had error in it. That 99.5% might in fact be 98%. In the Kelly formula, a deviation of that 1.5 percentage points would turn the "mathematical optimum" into an "actual loss" — any overconfidence is magnified mercilessly by geometric compounding.

Second: he was managing other people's money — pensions, retirement accounts. If some bond really did default, he could not leave his clients facing a situation where "a single bond dragged down the portfolio." Clients would redeem, the fund would be wound up — this is the greatest way an institutional investor dies.

Third: the "tail risk" of the bond market is far fatter than the surface data makes it look — once a recession or an industry shock hits, defaults erupt in clusters (the 2008 subprime crisis is the classic case). The "default probabilities" in the 1990s textbooks failed completely in 2008.

So he applied a very aggressive discount to the Kelly formula — cutting the mathematical "theoretical optimum" down to a 2% "practical ceiling."

2% amounts to "a very small fraction of Kelly" — extremely conservative, but enough to carry him alive through every crisis over forty-four years.

This is the highest use of the Kelly formula in the bond market — not treating it as an instruction for "how much you should bet," but treating it as a red line for "how much you must not exceed."

6.7 The Shock of a Few Numbers

Let us look at a few sets of numbers, to feel the power of "the 2% rule plus long-term compounding":

PIMCO Total Return Fund performance (1987-2014): - Cumulative net return of about 2200% (nearly 23 times) - In that period no single year saw the net asset value fall more than 5% - In that period no single bond caused the fund a daily drop of more than 0.5%

Bill's own wealth: - When he left PIMCO in 2014, his personal net worth was estimated at about 2.3 billion dollars - He gave away about 800 million of it — mainly to medical research and charity

The results for PIMCO's clients: - Tens of millions of American households held bonds through PIMCO - In twenty-seven years, the PIMCO Total Return Fund never in any single year lost clients more than 5% overall - This is a more stable wealth instrument than 90% of American stock funds — all of this was built on the "2% rule."

And this 2% rule came directly from the experience of a twenty-two-year-old unemployed youth playing four months of blackjack in Las Vegas.

6.8 A Cautionary Lesson: When Bill Abandoned PIMCO

Every story must have its cautionary lesson.

In September 2014, Bill left PIMCO — a dramatic exit, with complicated causes involving an internal power struggle that the outside world never fully understood.

He joined a smaller fund company called Janus Capital and launched a new fund called the Janus Henderson Global Unconstrained Bond Fund — the very name "Unconstrained" foretold something.

What does "unconstrained" mean? It means this fund was not subject to the usual constraints of bond investing — it could concentrate its bets, could use high leverage, could place directional bets on interest rates — unlike the PIMCO Total Return Fund, which had to diversify, had to keep leverage low, had to avoid directional wagers.

Put simply: Bill abandoned the 2% rule.

He began to place big directional bets — he believed European bond yields would rise (and so shorted German government bonds), and that U.S. inflation would restart (and so shorted U.S. Treasuries).

These bets were not necessarily wrong in direction — but the positions were too concentrated. He changed the 2% rule into bets of 10%, even 20%, per position.

The result:

• During 2015-2018, he shorted German bonds — but German bond yields kept falling (even falling into negative rates). He accumulated hundreds of millions of dollars in losses on his German-bond shorts. • He shorted U.S. Treasuries — but Treasury yields kept falling amid the 2019 wave of rate cuts. Again, hundreds of millions in losses.

In the few years Bill managed it, the Janus Unconstrained Bond Fund's net return was markedly below that of comparable funds, and far below the standard of Bill's PIMCO years.

In February 2019, Bill announced his retirement from Janus — an even more ignominious exit than his departure from PIMCO.

The lesson of this affair is extremely important: The discipline of the Kelly formula is a "skill." Once you abandon the discipline, your skill disintegrates along with it.

Bill's success in his PIMCO years came 70% from discipline (the 2% rule, diversified holdings, conservative leverage), and 30% from judgment (recognizing which bonds were undervalued).

When he got to Janus and abandoned the discipline, his "judgment" — however clever the judgment — could not save him.

This is not merely the story of one man, Bill. It is the shared warning to all users of Kelly:

What lets you make money is not your IQ, it is your discipline.

6.9 Bill's Reflections in Later Years

Bill turned eighty in 2024 and lives in California. In his memoir and his interviews he has reflected many times:

"My success at PIMCO was 95% because I knew how to control risk. Judgment accounted for only 5%.

The most important thing I learned in blackjack was not 'how to win money,' it was 'how not to lose it all.'

With any bet you place, you must ask yourself: 'If this bet is completely wrong, what happens to me?' If the answer is 'it would cost me 50% of my bankroll' — don't bet, no matter how much edge this bet appears to have.

Later I went to Janus, and I forgot this rule. I thought I had already 'made my name and my fortune' — and the market took three years to teach me the price of forgetting the rule."

This is the true spirit of the Kelly formula —

It is not about "how to win." It is about "how not to be eliminated."

And so long as you are not eliminated, compounding will do the rest for you.

Chapter Summary: The Three Things Bill Gross Left Behind

First: the Kelly formula's "2% rule" can be carried from the casino into the bond market — this rule still held on a scale of two trillion dollars under management.

Second: a conservative fraction of Kelly (about 1/5 Kelly) is far safer than full Kelly — especially when you are managing other people's money.

Third: discipline is skill. Once you abandon discipline, your judgment — however clever the judgment — cannot save you either (the cautionary lesson of Janus Unconstrained). In the next chapter, we will look at a man who never publicly used the Kelly formula at all — yet everyone who has seen his positions says: "This is Kelly."

He drinks Coca-Cola, loves cherry pie, and lives in a single-story house in Omaha.

His name is Warren Buffett.

Chapter 7 · The Oracle of Omaha's Sleight of Hand

7.1 A relationship no one can deny, yet no one can fully prove

1976. California.

Edward Thorp — whom you already met in Chapters 4 and 5 — was forty-four that year, at the height of PNP. One evening he attended a finance seminar and ran into an investor he had never met but whose reputation he knew well.

The man was two years younger than Thorp, forty-six, dressed in a rumpled suit, wearing black-framed glasses, drinking a Coca-Cola. He had just flown in from Omaha — that small city in Nebraska that hardly anyone ever visits.

His name was Warren Buffett. By then he was already one of the most famous investors in America — his partnership fund had returned 30% annualized from 1956 to 1969, and after shutting it down in 1969 he had begun using Berkshire Hathaway as his investment vehicle.

Buffett and Thorp talked all afternoon. About what? Bridge. Both were expert bridge players.

But in between the bridge, Buffett quietly probed Thorp's strategies — convertible-bond arbitrage, statistical arbitrage, the Kelly formula. And

Thorp quietly probed Buffett's — value investing, concentrated positions, long-term holding.

Thorp later wrote in his autobiography:

"By the end of the conversation I had a strange feeling. Warren had not read Kelly's paper. He had never studied the mathematical Kelly formula.

But everything he did — allocating a large fraction when he was convinced he had a significant edge, refusing to bet when there was no opportunity, always controlling the maximum loss on any single position — was the very same thing the Kelly formula was saying.

By intuition he had arrived at exactly the same conclusion as the mathematics."

This is the theme of this chapter — that relationship between Buffett and the Kelly formula which no one can deny, yet no one can fully prove.

7.2 Buffett never mentions the Kelly formula in public

Here is the fact:

Warren Buffett, in more than sixty years of investing, has almost never publicly mentioned the Kelly formula.

He has written dozens of letters to shareholders. He has given hundreds of interviews. He has published books, appeared on magazine covers, been studied countless times. Nowhere does he explicitly say, "I use the Kelly formula."

But —

Charlie Munger — Buffett's partner, vice chairman of Berkshire — has publicly acknowledged the influence of the Kelly formula many times.

In a 1997 speech Munger said:

"A highly diversified portfolio needs only four stocks. Smart people bet heavily when the world offers them the opportunity. That is what Kelly is telling you."

In an interview in the 2000s Munger said:

"The mathematical rationale for concentrated investing can be traced back to Kelly and Shannon and that generation. We don't write down the formulas, but we act according to them."

Buffett himself almost never comments on such things. He seems to feel that reducing investing to "he used the Kelly formula" oversimplifies investing itself — investing also involves qualitative judgment, understanding of the business, trust in management, a feel for market sentiment. None of that is in the Kelly formula.

But the way he acts — is astonishingly consistent with what the Kelly formula recommends.

7.3 Case one: American Express, 1963 — the most classic "Kelly-style big bet"

November 1963.

A company called Allied Crude Vegetable Oil Refining was found to have perpetrated a massive salad-oil fraud — using a letter-of-credit mechanism, they had falsely reported hundreds of millions of pounds of salad oil that did not exist at all, using it as collateral to swindle loans.

That company's bank was American Express. American Express faced a potential loss of $150 million from the scheme (in 1963 dollars, roughly $1.5 billion today).

When the news broke, American Express's stock was cut in half — from $60 down to $35. The market concluded American Express was going to go bankrupt.

What did the thirty-two-year-old Warren Buffett — who was then still running Buffett Partnership Limited — do? He did one thing personally: he walked into several restaurants, banks, and shops in Omaha and asked the cashiers:

"Do you still accept American Express traveler's checks?" "Are your customers still using American Express cards?"

The answer was: still using them, still accepting them.

From this simple piece of "street research," Buffett reached a judgment:

The financial loss from the salad-oil scandal was real (about $150 million), but American Express's core business — traveler's checks and credit cards — was entirely undamaged.

The brand was intact. Customer trust was intact. Earning power was intact.

The market had overreacted.

He did some arithmetic:

• Upside probability p ≈ 80% (the brand moat was intact, a return of the stock price was highly likely) • Upside magnitude: about 3–5x (a rise from $35 back above $100 was reasonable) • Downside magnitude: about 30–50% (it had already fallen once; there was limited room to fall again)

Plugging into the "value-investing simplified version" of the Kelly formula (this simplified form was later worked out by Mohnish Pabrai; in essence it is (p × Upside − q × Downside) / Upside):

the theoretical Kelly position was about 78%

What did Buffett do?

He bought American Express with 40% of his partnership's capital — the maximum proportion the partnership agreement allowed.

A 40% position was already the most aggressive "concentrated position" of any partnership fund of the time. Many investors thought he was crazy — putting 40% of his capital on a "scandal company."

And in the years that followed?

• From 1964 to 1966, American Express's stock rose from $35 to $180 — more than 5x • Buffett's partnership fund made a cumulative profit of about $20 million on this one position — equivalent to more than 50% of its total assets under management

In the purchasing power of the time, $20 million was equivalent to roughly $200 million today.

More important still — American Express made Buffett's name explode. From that moment on, the amount of money he went on to manage began to grow exponentially.

7.4 The "Kelly legitimacy" of this one bet

Let us look again at the American Express bet in the language of the Kelly formula:

• Full-Kelly recommendation: 78% • Buffett's actual bet: 40% — roughly "1/2 Pabrai-Kelly"

Why did Buffett bet only 40% instead of 78%?

The reasons are very modern — the same reasons we saw earlier for Thorp and Bill Gross:

First: he had uncertainty about his estimate of the win rate. He intuitively judged it at 80%, but it might actually have been 70%, might have been 65%. Halving Kelly can absorb 10–15 percentage points of estimation error.

Second: the partnership agreement had a hard cap of 40%. This was an external constraint — but it happened to be close to the optimal fraction of "half-Kelly." External constraints are often another manifestation of the Kelly spirit.

Third: he needed to keep capital in reserve for other opportunities. Even if American Express was a good opportunity, there would be more good opportunities in the future. Putting everything on one bet means giving up future optionality — which violates another implicit principle of the Kelly formula.

So the 40% Buffett actually bet — was not mathematically optimal — but was his "battlefield Kelly." This is a use smarter than pure mathematics.

7.5 Case two: Coca-Cola, 1988 — the power of long-term compounding

July 1988.

Warren Buffett began buying Coca-Cola stock for Berkshire Hathaway — at an average of about $5.22 per share (adjusted for splits).

By the end of 1989 he held about 93.4 million shares — at a total cost of $1.299 billion. That was equivalent to 35% of Berkshire's stock portfolio at the time — again a major concentrated position.

Why Coca-Cola? Buffett's "Kelly logic" ran like this:

• Brand moat — Coca-Cola was a brand more than a hundred years old, almost impossible for a competitor to replace • Room for global expansion — in the 1980s Coca-Cola was only just beginning to enter China, India, Eastern Europe, and Africa • Earning power — ROE consistently above 30%, with steadily growing free cash flow • Reasonable valuation — the 1988 P/E was only about 13–15x, well below the historical average

Full-Kelly's "value-investing simplified version" recommended: a concentrated position of 40–50%. Berkshire's actual position: about 35% — once again "half-Kelly, leaning conservative."

The story that followed is well known —

By 2024, Berkshire still held the 400 million shares of Coca-Cola it had bought back then (the figure after multiple splits). Those shares had a market value of about $25 billion — a cumulative capital appreciation of about 19x.

Add the roughly $7.5 billion in dividends accumulated over 36 years — and the total return on this one investment exceeded 25x.

On an annualized basis, the compound return from 1988 to 2024 was about 10%. Doesn't sound especially extravagant? But this was $1.299 billion turning into $32.5 billion — 32x.

7.6 Case three: Apple, 2016 — a new story for the "old Buffett"

Many people think of Buffett as an "old-school" investor — buying only consumer goods, only banks, never touching technology.

Until 2016.

That year Berkshire began buying Apple stock on a large scale. By 2018, Apple at one point made up 40–45% of Berkshire's stock portfolio — one of the heaviest single positions in Berkshire's history.

What was Buffett's reasoning? At Berkshire's 2017 shareholders' meeting he said:

"I didn't invest in tech companies before, because I couldn't understand them. But Apple isn't a tech company — it's a consumer-goods company. Its product is the iPhone, but its real moat is the iOS ecosystem, it's user stickiness, it's the brand — and those are no different from Coca-Cola.

Once I saw that clearly, I dared to bet big."

By 2024, Berkshire's Apple position had earned a cumulative profit of more than $130 billion — the most profitable single investment in Berkshire's history.

— Once again a textbook "Kelly-style big bet": high conviction + large fraction + long-term holding.

7.7 Munger's "four stocks are enough"

If Buffett's attitude toward the Kelly formula is "use it in secret, don't comment in public," then Charlie Munger is the one who brought the Kelly spirit out into the open.

Munger — born in 1924, trained as a lawyer, who began working with Buffett in 1959 — was Buffett's lifelong partner and "thinking partner." In his public speeches he repeatedly stressed a few principles, each of them a non-mathematical version of the Kelly formula:

Principle one: "Concentrated investing is smarter than diversification"

"Academia tells you that you should hold 50 stocks to spread the risk. That's meant for people who don't know what they're doing.

If you know what you're doing — you've seen clearly the value of 4 stocks — then you should put your money in those 4.

Mathematically, this is what the Kelly formula tells you: high conviction = big position."

Principle two: "Wait for the best opportunity"

"Most of the time, there are no truly good opportunities in the market. What should you do then? Nothing at all.

But when the opportunity really comes — you strike hard.

The smartest thing Buffett did wasn't the few big bets he made; it was that he waited 5 years before making those few big bets.

In the Kelly formula this is called not making a negative-expectation bet. In life, it's called patience."

Principle three: "Invert"

"If you want to know 'how should I invest' — don't ask that question.

Ask: 'How can I avoid investment failure?'

Then turn it around — avoid the things that would make you fail, and what's left is good investing.

This is the core of the Kelly formula: it first tells you not to go broke, then tells you how to make money."

These views of Munger's — though he never writes down a formula — are in essence the human expression of the Kelly formula.

7.8 A dissenting voice: Mohnish Pabrai's change of heart

But the story cannot be told from only one side.

Mohnish Pabrai — an Indian-American value investor, hedge-fund manager, and famous follower of Buffett and Munger — formally introduced the Kelly formula to the value-investing world in his 2007 book The Dhandho Investor.

In the book Pabrai gave a "value-investing version of Kelly": F = Edge / Odds

and recommended that value investors use this formula to decide the position size of a single stock.

The book was a bestseller. Many young value investors, following Pabrai's method, put heavy positions in a single stock — say, some cheaply valued small-cap, theoretical F = 30%, and they would stake 30% of their capital on it.

The result?

Many of them lost badly.

Why? Because they overestimated their own edge — in value investing, the estimate of "intrinsic value" is subjective, highly dependent on the estimator's judgment. Not every stock trading at a P/E of 5 is undervalued — many really do have problems, and the market is right.

The Kelly formula accepts only true probabilities, not subjective estimates.

Pabrai himself later publicly admitted this error. In several interviews he said:

"Introducing the Kelly formula to the value-investing world was a mistake.

The Kelly formula assumes your estimate of the win rate is true, objective, quantifiable — that's how it is in blackjack and at the racetrack.

Value investing is not like that.

The 'win rate' in value investing is a subjective probability you estimate yourself. When you overestimate it — and you almost certainly will overestimate it — the Kelly formula pushes you toward dangerous over-betting.

The true version of value investing is 'many small bets + extreme conservatism' — not the Kelly formula."

This is a deep warning from the Kelly formula:

It works only in scenarios where you can objectively quantify the win rate.

When the win rate is a subjective estimate — you need to be far more conservative than Kelly.

Buffett's success lies not in his using Kelly — but in his applying Kelly's "bet big only on high win rates" principle to an extremely demanding standard: he bets big only on things he is 95% sure of.

And most of the time, he believes the things he is 95% sure of are very few — so he always holds a large amount of cash (Berkshire's cash has long been 15–30% of total assets).

This is Buffett's true "Kelly" — being conservative to the extreme by overestimating his own uncertainty.

7.9 A set of numbers

Let us end this chapter with a few sets of numbers:

Buffett's position-concentration statistics (1965–2024): - The top 5 holdings averaged 65–75% of Berkshire's stock portfolio - The top 10 holdings averaged 80–90% - The historical peak weight of the single largest holding: about 45% (Apple, 2018–2020)

Berkshire's cash reserves: - Long-term 15–30% of total assets - Reached a historical high of over $300 billion in 2024 - This cash is not "lazy money" — it is ammunition prepared for "enormous opportunities"

Berkshire's returns: - Cumulative return of about 5,000,000% (50,000x) from 1965 to 2024 - Annualized compound of about 20% - The S&P 500 over the same period was about 10% annualized

50,000x — this is the compound result of the Kelly spirit over 60 years.

And all of it — in Buffett's and Munger's own words — springs from one simple principle:

Bet a large fraction when you have a real edge; wait patiently when you have none.

This principle, and that half-line John Kelly wrote down at Bell Labs in 1956, are the same thing.

Chapter summary: Buffett's "non-mathematical version" of the Kelly formula

First: the mathematical legitimacy of concentrated positions — Buffett repeatedly bet heavily with a 40% position, which is entirely reasonable under the "value-investing simplified version" of the Kelly formula.

Second: holding a large amount of cash is part of the Kelly spirit — don't bet when you have no edge; when the edge appears, strike a lethal blow.

Third: the Kelly formula works only when you can objectively quantify the win rate — the "subjective win rate" in value investing leads people into overconfidence. Buffett's "response" is to raise his own standard for the win rate to 95%+ — most of the time, he simply doesn't bet at all.

Fourth: discipline always matters more than intelligence. Bill Gates once asked Buffett: "What is the single most important word for your success?" Buffett said: "Focus." — This is the deepest expression of the Kelly spirit.

In the next chapter, we will look at a man you may never have heard of, but who may be the person with the highest single-investment profit in history — using one computer and one Kelly formula, he made off with a billion dollars at the Hong Kong racetrack.

His name is Bill Benter.

Chapter 8 · The Billion-HKD Horse-Racing Algorithm

8.1 A Master of Physics Barred from the Casinos

Early 1984. Hong Kong. Kai Tak Airport.

A plane from Las Vegas touches down. Two Americans step off—

• A tall young man with a full beard and a rumpled shirt, twenty-nine years old, a background in a master's degree in physics, a weariness in his eyes that says, "I've just been swept out the door of every casino in America." • An Australian of about sixty, lean and tall, silent, wearing a hat.

Their names are Bill Benter and Alan Woods.

The passport pages in their hands are stamped over and over with "entry prohibited"—80% of the large casinos in America have put them on a blacklist. Walk into any American casino, and the moment the front desk checks their ID, security politely shows them the door.

They have come to Hong Kong not as tourists. They have come to bet on the horses.

And the weapon in their hands is the same one Thorp carried when he walked into Reno in 1961—

a computer + a probability model + the Kelly formula.

The difference is that what Thorp conquered was blackjack; what they meant to conquer was the deepest, most complex, largest horse-racing market in the world—the Hong Kong Jockey Club (HKJC).

Over the next seventeen years—from 1984 to 2001—they took more than one billion US dollars out of the Hong Kong racetrack.

By the official recognition of the 2025 Guinness World Records: Bill Benter is the most profitable player in the history of any single wagering game.

8.2 Bill's Road to Growing Up

Bill Benter was born in 1957 in Pittsburgh, Pennsylvania—a small steel-industry town. His father was a second-generation German immigrant, his mother an Irish-descended housewife. The family was of ordinary means.

From childhood he was good at math and good at physics, but he liked even more "using mathematics to solve strange problems"—in high school, for instance, he was obsessed for a time with chess, obsessed with bridge, and studied how to analyze these games with probability theory.

In 1977 he earned a bachelor's degree in physics from Case Western Reserve—by plan he was to go on to a doctorate in physics and pursue academic research. But in the summer he was about to graduate, he did something that turned his life completely around—

he read Edward Thorp's Beat the Dealer.

That book drove him wild. He canceled his doctoral plans and flew straight to Las Vegas, took $8,000 as a bankroll, and began playing blackjack by Thorp's method.

He met Alan Woods—an Australian veteran with fifteen years of professional gambling experience—in Las Vegas. The two formed a "professional blackjack team," and using card counting + Kelly betting, they roamed the casinos across America for six years.

But the casinos learned to fight back:

• In the early 1980s, all the large casinos installed prototype facial-recognition systems • Photographs of card counters were circulated to casinos across America and Canada • Walk into any casino, and within five minutes you would be shown out

By the end of 1983, Bill and Alan had been blacklisted by 80% of America's casinos. Their "careers" were, in effect, over.

They sat in a Las Vegas café, watching the December 1983 snow. Alan said:

"There's nothing left to play in America. Hong Kong has horse racing."

8.3 Why Hong Kong Racing Is a Special Place

Hong Kong racing has several distinctive features that make it "a paradise for the professional gambler":

Feature one: the pari-mutuel system

In Hong Kong you are not betting against a bookmaker—all the wagers pour into a single pool, the HKJC skims off a certain percentage as a "management fee" (about 17.5%), and the rest is distributed to the winners in proportion to their bets.

This means the odds are not set by a bookmaker; they are the collective decision of all the bettors. The bet you place affects the odds—the more you bet, the lower the odds.

It also means: if your probability model is more accurate than the "market consensus," you must win in the long run. Because you will find those horses that are "undervalued by the market"—and the odds on those horses happen to be higher than their true win probability.

Feature two: the largest racing pool in the world The total wagering on each race day in Hong Kong often exceeds HK$1–1.5 billion—roughly US$150–200 million. This is the deepest, most liquid horse-racing market in the world. A "professional syndicate," even betting large, will not significantly disturb the odds.

Feature three: the race data is extraordinarily complete

The HKJC publicly releases hundreds of data points for every horse in every race—bloodline, trainer, jockey, track, weather, pace, past performance, weight changes… This was, in 1984, the racetrack in the world best suited to "quantitative analysis."

Feature four: Hong Kong racing's takeout rate (17.5%), though not low, is among the lowest of the world's major racetracks—the takeout rate at some tracks in certain US states runs as high as 25–30%, making profit outright impossible.

Bill and Alan sized up these features. They flew to Hong Kong carrying a bankroll of $150,000—saved up through gambling and help from family.

8.4 The First Three Years: A Total Wipeout

1984–1986. They failed utterly.

Their model at the outset was simple—it used just a dozen-odd variables (speed, jockey win rate, track configuration) to build a simple regression. The result—

in three years they lost the entire $150,000.

Alan pulled out. In 1986 Alan told Bill: "I've had enough. I'm going back to Australia."—he later went on playing at Asian racetracks on his own, and in the end became a millionaire too (though not on Bill's outrageous scale).

Bill stayed on in Hong Kong alone. The money was gone. He scraped by earning a little money as an "analyst" for other professional gamblers at the HKJC—writing analysis reports for others, earning HK$200–500 apiece.

But he did not give up.

He did one crucial thing—he rewrote the entire model.

The new model used more than 70 variables—later expanded to more than 120:

• The horse's speed ratings over its last 5 races (not just placings, but the specific speed data of each race) • Days since its last race • The jockey's win rate over the past year (tabulated separately by track, distance, and horse type) • The trainer's win rate over the past year • Track weather (dry / wet / soft) • The starting gate (differences in "luck" between gates) • Weight change (added weight usually favors the horse) • Pace prediction (front-runner / mid-pack / closer) • The horse's trajectory of change over the past 12 months • Sire genetic traits • The trainer's overall recent form

He used a statistical model called multinomial logistic regression, letting the computer learn from historical data the relationships between these variables and "win probability."

This model looks very simple today—any junior data scientist in 2026 could write it. But in 1986–1987, Bill was working on an IBM XT personal computer (that's an 8088 processor, 640KB of memory)—it took him several months just to get it running.

8.5 The First Big Money of 1987

1987. Bill restarted.

The new model + the Kelly formula. He borrowed a few tens of thousands of dollars as a bankroll.

That year, he made $600,000—his bankroll multiplied dozens of times over.

In 1988, he made $3 million. In 1989, he made $6 million. Through the 1990s, his annual profit held steady between $10 million and $50 million.

By 2001—his seventeenth year in Hong Kong—his cumulative profit exceeded $1 billion.

8.6 The Subtle Application of the Kelly Formula in Horse Racing

Bill's use of "Kelly" was more complex than in blackjack. Let us look at his system:

Step one: compute each horse's true win probability

Model input: the 120-plus features of the dozen-odd horses in that race. Model output: each horse's true win probability p_1, p_2, …, p_{12}.

Step two: find the horses the market has undervalued For each race, the HKJC displays market odds o_1, o_2, …, o_{12}. Each odds figure corresponds to a "market-implied probability" $1/(o_i + 1)$.

If Bill's model says "$p_i >$ the market-implied probability," then this horse is undervalued—he should bet on it.

Step three: allocate capital with the n-way Kelly formula

The n-way Kelly is not simply "applying Kelly independently to each undervalued horse"—that would overstate the total position. It is a convex optimization problem:

$$\max_{f_1,\dots,f_n} \sum_{i=1}^{n} p_i \cdot \log\left(1 + b_i f_i - \sum_{j \neq i} f_j\right)$$

This requires solving a mathematical optimization problem. Bill's team ran this optimization on the IBM XT—at 1987's computing power, each race took several minutes to compute.

Step four: bet in batches

If, for a given race, Bill's "theoretical Kelly recommendation" is to bet HK$500,000—he will not lay it all down at once.

Why? Because:

• Betting too much at once will drive the odds down—under the pari-mutuel system your bet changes the odds • Drawing the HKJC's attention will get you put under surveillance • In case the signal is false (the model occasionally errs), batching lets you cut losses in time

His team would lay it down in 10–15 separate bets in the 1–2 hours before the race—each of a few tens of thousands of HK dollars, disguised as the behavior of "ordinary gamblers."

Step five: use "half-Kelly" or lower

Even when theoretical Kelly said to bet 5%, Bill would actually bet only 2–3%—once again the classic "fractional Kelly." The reasons are ones you already know: estimation error + liquidity + psychology.

8.7 A Detail That Seems Unbelievable

Many readers, reaching this point, will ask: "If Bill's model was so strong, why didn't the HKJC ban his account?"

The answer has several layers:

First: the HKJC didn't care. The HKJC is a nonprofit institution—after deducting the entire takeout (17.5%), it donates most of it to charity and education in Hong Kong. Bill made $50 million a year—the total he wagered may have been several billion—and the HKJC skimmed a "management fee" from his bets far exceeding that figure.

The money he made for the HKJC was more than the money he made for himself. Of course the HKJC would not drive him away.

Second: Bill's betting made the market more efficient.

When Bill bet on an "undervalued" horse, the odds would be pushed down—that is, closer to the true probability. This made the bets of other gamblers more accurate and raised the efficiency of the whole market.

The HKJC in fact welcomed "professional syndicates" like Bill's—they made the market smarter and more stable.

Third: Bill obeyed all the rules.

He never manipulated the odds, never bribed jockeys, never used inside information. All his bets were based on public data—he simply analyzed it more cleverly than others.

This is the cleanest version of applying the Kelly formula—making money purely on an information edge, relying on no improper means whatsoever.

8.8 Retirement in 2001

In 2001, at forty-four, Bill retired of his own accord.

Why? The reason is the same as Thorp's for closing Ridgeline in 2002—

"the market got crowded, and the edge was vanishing."

Starting in the late 1990s, more professional syndicates entered Hong Kong—Japanese syndicates, Singaporean syndicates, British syndicates. They too used machine-learning models, they too used the Kelly formula.

Bill's edge slid from the early 5–7% (far above market consensus) slowly down to 2–3%—still profitable, but the risk/reward ratio was no longer what it had been.

He made a clear-eyed decision: get out while the edge was still there.

He left Hong Kong with a fortune of about $1 billion and returned to Pittsburgh—the place where he was born.

His life now is very quiet. No luxury car, no private jet, no fame. He lives in an ordinary house on the outskirts of Pittsburgh.

What has he done? He has given away a great deal of money:

• In 2007 he founded the Benter Foundation, focused on local Pittsburgh education and the arts • He funds several projects at the University of

Pittsburgh • He funds a number of cultural institutions and museums in Pittsburgh

He also did something that moved the mathematical world—he established the "Benter Prize in Applied Mathematics," a prize of $50,000 awarded each year to a researcher who has made outstanding contributions to applied mathematics.

This is his way of giving back—to the Kelly formula, to Thorp, to Shannon, to the intellectual tradition that let him live an unbelievable life through mathematics.

8.9 Some Quotations from Bill's Interviews

After retiring, Bill gave a few interviews. Several passages he spoke are especially worth remembering:

On the Kelly formula:

"The Kelly formula is not about how much you win, it's about not getting eliminated.

The hardest thing in my career was not building the model—it was facing, several times a day, the temptation of the model telling me 'you should bet HK$500,000 on this race.'

True discipline is betting HK$500,000 for real when HK$500,000 sounds like too much; and betting only HK$500,000 for real when HK$500,000 sounds like not enough.

Most people cannot manage the second one."

On luck:

"I'm often asked: 'How much of your success is luck?'

My answer: the overwhelming majority of it is luck.

Hong Kong racing in 1984 had the special system of a takeout of only 17.5%—that was luck. Alan and I were driven out of Las Vegas and only then turned to Hong Kong—that was luck. When I nearly gave up in 1986, I chose to go on—that was luck (any rational person at the time would have quit). Hong Kong's racing data in the 1990s was especially complete—that was luck.

What I did was simply amplify that luck—through disciplined Kelly betting."

On failure:

"In my first three years I lost my entire $150,000. Those were the three most important years of my life.

If I had made a billion in my first three years, I would have overrated myself. I would have thought it was my cleverness. I would have bet bigger. I would have gone bankrupt. It was the failure of those three years that taught me how deep 'I don't know' really goes.

Everyone who can truly use the Kelly formula well has first been through the process of wiping themselves out—it's a required course."

8.10 A Set of Numbers That Quiet You Down

Finally, a set of numbers.

Bill Benter's career statistics (1984–2001):

• Initial bankroll (including borrowing): about $150,000—all lost in the first three years • Initial bankroll after restarting: about a few tens of thousands of dollars • Cumulative profit: about $1 billion (recognized by Guinness World Records) • Maximum drawdown over the period (at the team level): about 30% • Total number of bets: an estimated 500,000–1,000,000 (cumulative over 17 years)

His model's win rate (vs. market consensus):

• Model IC (information coefficient): about 0.12–0.15 • Market-consensus IC: about 0.10 • Gap: about 2–5 percentage points

It was these 2–5 percentage points of "information edge"—through hundreds of thousands of repetitions over 17 years + the sensible amplification of the Kelly formula—that accumulated into $1 billion.

This is the purest, most romantic victory in the seventy years of the Kelly formula.

Chapter Summary: Three Lessons from Bill Benter

First: the edge need not be large—Bill's edge was only 2–5 percentage points above market consensus. The key is to amplify that tiny edge sensibly through the Kelly formula, and then to let the law of large numbers take effect through hundreds of thousands of repetitions.

Second: early failure is a required course—Bill lost his entire bankroll in his first three years. Had he made big money at the start, he would have overrated himself and, in the end, gone bankrupt. Only after being wiped out do you earn the right to talk about Kelly.

Third: get out before the edge vanishes—Bill retired of his own accord in 2001, Thorp closed down Ridgeline of his own accord in 2002. Every edge is temporary, and the ability to recognize this clear-eyed is rarer than the edge itself. In the next chapter, we will look at two figures who leave the American sports-betting world both admiring and exasperated—each using a completely different method (one making prop bets off "coaches' behavioral patterns," the other on 30 years of discipline + a Kelly portfolio across many games), and each taking hundreds of millions of dollars out of sports betting.

They are Bill Walters and Haralabos Voulgaris.

Chapter 9 · The Two Gamblers Who Gave the House a Headache

9.1 Why Sports Betting Is the Kelly Formula's Natural Battlefield

Horse racing is pari-mutuel — you bet against all the other bettors. Sports betting usually follows the bookmaker model — you bet against the house.

What is the difference?

In a pari-mutuel system (Bill Benter's Hong Kong horse racing), what you are looking for is "a horse the market consensus has undervalued." In the bookmaker model (the NBA, the NFL, the Premier League), what you are looking for is "a game the house has mispriced."

The house makes mistakes too. The house is a commercial enterprise; it wants to earn its 4–5% cut (the vig), and what it cares about most is a "balanced book" — as long as the amounts wagered on both sides are close, it pockets the vig no matter what, indifferent to the outcome itself.

This means the odds the house sets reflect not the "true probability" but the "odds that balance the money on both sides." If the public tilts irrationally to one side, the house's odds drift away from the true probability — leaving an edge for the professional gambler.

This is the fundamental reason "professional players" in American sports betting have been able to make money since the 1980s.

And the two protagonists of this chapter, using completely different methods, each walked away with over a hundred million dollars from this market.

9.2 The First: Bill Walters, "the Warren Buffett of Sports Betting"

Bill Walters — nearly everyone in today's American sports-betting world has heard the name.

He was born in 1946 in Kentucky, into poverty — his father died when he was eighteen months old, his mother was an alcoholic, and he was raised from childhood by his grandmother.

But from an early age he had an astonishing intuition for numbers. He started playing pool at nine, and at eleven he was already winning money from grown men in the pool halls. He dropped out of school at thirteen. By the time he was twenty, he had already played thousands of games of poker and thousands of games of pool in Kentucky's assorted little casinos and bars, and had made more than half a million dollars — richer than the great majority of the local businessmen.

In the 1970s he turned to sports betting — American football (the NFL) and basketball (the NBA, the NCAA). He built his models by hand — memorizing every game of every team — and slowly learned to "spot the house's mispricings."

By the 1980s he had become one of the largest sports-betting organizers in America. He built a team called "The Computer Group" — more than ten people, including statisticians, mathematicians, and former sports professionals — who together built the models, analyzed the games, and decided the bets.

Over the next thirty-plus years, Walters almost never had a losing year.

9.3 Walters's Kelly Philosophy

Walters does not write formulas — he is not a man with a mathematical background. But his hands-on method conforms entirely to the spirit of the Kelly formula:

First: never go all-in on a single game

Walters said again and again in interviews:

"Any single bet larger than 5% of my account, I won't place. I would rather miss the opportunity than let any single game be able to ruin me."

That 5% is roughly equivalent to "quarter-Kelly" — a very conservative fractional Kelly.

Second: bet on dozens or even hundreds of games at once

Walters's team bet on 100-plus games a week — American football, basketball, baseball, rugby, tennis, nearly every major sport.

Why? Because:

• The "signal-to-noise ratio" of a single game is very low — any one of them can lose • Stack 100 games together, and the law of large numbers takes hold — the steady return of a long-term edge surfaces

Third: actively work the market

Walters had another move — his team deliberately placed bets in batches, luring the public to follow.

How exactly? He would first place a small "exposure bet" — letting his wager be seen by the market's "followers" (the retail bettors who tail the professionals). The retail crowd follows in, and the odds begin to shift. Then he loads up in the opposite direction — when the odds have drifted to an extreme, he places a large bet, eating up the extra edge from this "odds drift."

This is a game only a large professional group can play — it requires the scale of capital, the market's attention, and a sense of timing. In essence, it turns "the market's perceptions" into another source of edge.

Fourth: strict financial discipline

Walters's organization was run like a hedge fund — with a CFO, risk control, and audits. Every single bet was recorded, reviewed, and dissected.

In 2017 he was charged with insider trading and imprisoned (not for anything to do with sports betting, but for a separate stock insider-trading case — he had used inside information from a golfing buddy to trade stocks), but even in prison his sports-betting organization kept running — because his "system" had already been institutionalized and did not depend on him alone.

In 2024 he was released and reappeared on the Las Vegas sports-betting scene. He published a memoir, Gambler, in which he set out in detail the sports-betting philosophy of his thirty-plus years.

Estimated cumulative profit: several hundred million dollars.

9.4 The Second: Haralabos Voulgaris, the God of NBA Betting

Voulgaris is a completely different man.

He was born in 1975 in Winnipeg, Canada — into a family of Greek descent, with a father who was a heavy gambler and who gambled away the family's fortune. Voulgaris lived through his family's bankruptcy as a child — he holds a deep-rooted, complicated feeling about gambling: on the one hand he hates it, on the other he is fascinated by it.

He attended a few years of university without finishing, and began betting on the NBA.

What he attacked was not the conventional "team win/loss lines" — he chose a particularly small, particularly specialized niche market: the NBA prop bet (proposition bet).

What is a prop bet? For example:

• "Will this player score more than 30 points in this game?" • "With his team up 15 with 5 minutes left, will the coach pull his starters?" • "In this game, which half will have the higher total score?"

The pricing of these prop bets is far cruder than the "win/loss lines" — because the house invests less effort in these niche markets. Voulgaris found several niche categories in which the house was systematically mispricing. The most famous was the "first-half/second-half total-score gap" that he cracked —

He discovered that the house set the first-half total at exactly half the full-game total. But through statistical analysis, he found that for certain teams and in certain situations, the scoring patterns of the first and second halves differed significantly — the first half was often a slow start (the defensive phase), and the second half accelerated (defense collapsing under fatigue).

He modeled and quantified this "half-time asymmetry," and found a class of prop bet with an edge of roughly 5–8% — a very large edge.

Over the next several years, he placed prop bets on every NBA game — on average 30–50 a day (the NBA has 5–10 games a day, and each game has several prop bets).

The peak-period numbers are staggering:

• Total daily wagers: close to a million dollars • Win rate: about 60–70% • Annualized profit: an estimated 15–30 million dollars a year • Cumulative profit 2000–2014: about 150 million dollars

9.5 Voulgaris's Kelly System

Voulgaris is a strict disciple of Kelly — even stricter than Walters.

He described his "working day" in an interview:

"I get up at 6 every morning. The first thing is to review last night's games — what my model said, what I bet, what the result was, where the model was wrong.

At 9 I start modeling tonight's games. One hour per game — reading every injury report, lineup change, coaching tactical analysis, and the past 10 head-to-head matchups.

At 3 in the afternoon I feed the model's results into my 'Kelly calculator.' Given the edge, given the variance, given my current account, how much should I bet?

Cap for a single game: 2% of the account. Cap for stacking several games in one night: 8% of the account.

Then I place the bets — spread across 15–20 different betting accounts — to keep any single house from spotting me.

At 11 at night the games tip off, and I watch every one of them live. When they're over, I record all the outcomes. Sleep.

At 6 the next day it starts again.

This was my life for 12 years."

— This is the real life of a professional-grade Kelly user: dull, disciplined, long hours, unromantic.

9.6 Voulgaris's Crisis: the Edge Disappears

But Voulgaris's story is not only about "how much he won" — more important is how he handled the edge disappearing.

Around 2014, he found that the edge of his "first-half/second-half asymmetry" model was steadily declining — from an early 5–8% down to 2%, then down to 1%.

Why? Because:

• The house learned — they hired quant analysts and folded the logic of "half-time asymmetry" into their pricing • Other professional gamblers copied his method • Data-driven sports-betting tools spread — even amateur players could now buy statistical software

By 2014, his model was essentially unable to turn a profit.

Worse still, he himself made a fatal mistake —

During 2010–2014, he violated his own Kelly discipline. Firmly convinced that "his own judgment was still valid," he kept placing large bets and failed to scale down his position as the edge declined.

The result — over 2014–2016 he lost tens of millions of dollars cumulatively — nearly wiping out the profits of the preceding years.

9.7 Voulgaris's Rebirth

In 2016, Voulgaris admitted defeat. He did something professional gamblers rarely do —

He stopped entirely and started learning again.

He hired a machine-learning team and rewrote his whole modeling framework from scratch — this time not just NBA prop bets, but comprehensive NBA game prediction. He gave this new model a name, "Ewing" — after the NBA star Patrick Ewing.

The Ewing model used:

• Players' real-time performance data (ball speed, distance moved, position) • Team chemistry metrics (how well the players meshed with one another) • Coaches' decision patterns (timeout timing, substitution patterns) • Psychological indicators (state changes after consecutive losses)

The new model's edge climbed back to 3–5% — profitable. Over 2017–2019, Voulgaris rebuilt his account — accumulating tens of millions of dollars again.

In 2019 something happened that changed his life —

The Dallas Mavericks (an NBA team owned by Mark Cuban) came to him and invited him to be the team's "analytics chief" (Director of Quantitative Research and Development).

Cuban's reasoning: "Voulgaris is the most accurate predictor of player and coach behavior in NBA history. No one understands this league's micro-signals better than he does."

Voulgaris accepted the position. He went from "betting against NBA teams" — to "working for an NBA team."

This symbolically ended his career as a "professional gambler." But the methodology he left behind — through his interviews on ESPN, through his work with the team — has become one of the standards of NBA quantitative analysis.

9.8 The Two Different Lessons Walters and Voulgaris Give the Kelly Formula

Let us compare the two:

Bill Walters Haralabos Voulgaris Origin Kentucky poverty Greek descent, Winnipeg, Canada Education Dropped out at 13 Did not finish university Style Large organization, Individual + small team, multi-person team, focused on the NBA across many sports Edge source House mispricing + NBA prop bet micro-modeling retail-following effect Position Single bet 5%, Single bet 2%, rules diversified across 100+ games 8% per night Cumulative Several hundred million About 150 million dollars profit dollars

Fate Imprisoned (insider Became an NBA team analyst trading), continued after release

What they have in common — and these are the most core tenets of the Kelly spirit:

First: a strict cap per bet (Walters 5%, Voulgaris 2%). This is the "discounted version" of the Kelly formula — conservative down to one-quarter or one-eighth Kelly.

Second: letting the law of large numbers take hold through a large number of repetitions. Both men bet on 100-plus games a week — rather than going all-in on a single game.

Third: staying alert to edge decay — Voulgaris failed to adjust in time in 2014 and paid the price of tens of millions; his "rebirth" came because he admitted the edge was gone and rebuilt his model. Fourth: discipline matters more than cleverness — neither of them was the "smartest" man, but their discipline let them survive to the day the law of large numbers took hold.

9.9 A Cautionary Lesson: Ignoring Correlation

Finally, a cautionary lesson — the biggest mistake amateur players make when using the Kelly formula.

The "independent form" of the Kelly formula assumes each bet is independent — that there is no correlation among their outcomes.

But in sports betting, multiple games are often highly correlated:

• Eight NBA games on the same night — if the total-score lines are broadly set too low (referee style, team fatigue), all eight go "under" • The "momentum effect" within the same league — a certain tactical style suddenly works, and the whole league changes at once • A uniform adjustment by the platform/house — one house fixes a bug, and all their previous "errors" suddenly vanish

If you use independent Kelly to size each of eight games separately, each might be recommended at 2% — 16% in total. But these eight are actually correlated — the real exposure may be equivalent to 6–8% of a single game.

The correct approach is to use "multi-game Kelly" — bringing the covariance matrix into account:

$$f^* = C^{-1} (b \cdot p - q)$$

But amateur players almost never do this — they use independent Kelly, and the result is systematic over-betting and long-term losses.

This is one of Walters's and Voulgaris's most important advantages — both of them handled correlation rigorously: total exposure in a single night never exceeded 8%, and multiple games were adjusted for correlation among them.

Chapter Summary: The Kelly Spirit in Sports Betting

First: the edge comes from "house mispricing" + "irrational public following" — these two edges gave professional players a market worth hundreds of millions of dollars a year from the 1980s to the 2020s.

Second: a per-game cap + multi-game diversification + correlation adjustment — this is the "three-piece set" of sports-betting Kelly. Skip any one of them, and you lose in the long run.

Third: the edge will disappear — the lesson of Voulgaris's failure in 2014–2016: when the edge is declining, failing to reduce your position in time will quickly give back your earlier profits. Fourth: discipline is scarcer than IQ — Walters and Voulgaris were not the smartest men, but their discipline over thirty years / fifteen years let them survive.

In the next chapter, we will look at a completely opposite world — one with no discipline, extreme leverage, and a market that never stops, 24 hours a day.

That is cryptocurrency.

What we will look at is the 72 hours in May 2022 when LUNA fell from 80 dollars to 0.0001, FTX going from the world's second-largest exchange to an empty shell within a week in November 2022, and the stories of countless ordinary investors liquidated to zero on perpetual futures contracts.

The protagonist of that chapter is not any single person — it is all those people who had never heard of the Kelly formula, or who had heard of it but chose to ignore it.

Chapter 10 · Wiped Out Overnight: The Kelly Penalty in Crypto Markets

10.1 A Story That Needs No Protagonist

Each of the preceding chapters had a protagonist—Kelly, Shannon, Thorp, Bill Gross, Buffett, Benter, Walters, Voulgaris.

This chapter is different.

The "protagonist" of this chapter is not any particular person—it is an era and a crowd of people:

Over the eight years from 2017 to 2025, the wealth that evaporated from the global cryptocurrency market is estimated at more than 5 trillion dollars. The overwhelming majority of that money—over 80%—came from ordinary retail investors.

They had never heard of the Kelly formula. Or they had heard of it but thought "that old, outdated thing is useless." Or they had heard of it but thought "I'm smarter than John Kelly was in 1956."

This chapter is for them—and for everyone who might repeat these stories in the future.

10.2 Why the Crypto Market Is "the Cautionary Counterexample to the Kelly Formula"

The crypto market has several features that make it the hardest place of all to maintain the discipline of the Kelly formula:

Feature One: Extreme Volatility

• The S&P 500 stock index has an annualized volatility of about 16-20% • BTC (Bitcoin) has an annualized volatility of about 60-80% • Small altcoins frequently have annualized volatility above 150%

In other words, crypto markets are 4 to 8 times as volatile as the stock market. This means the "theoretical position" the Kelly formula prescribes is usually a reasonable 1/4 to 1/8—and the vast majority of ordinary investors cannot be that conservative.

Feature Two: 24/7 Trading

Stock markets have opening and closing bells—while you sleep, the market does not trade. The crypto market runs 24 hours without pause—while you sleep, you may be liquidated.

This makes the "psychological pressure" far greater than in traditional markets.

Feature Three: 125x Leverage

The major crypto exchanges (Binance, OKX, Bybit) allow retail investors to use up to 125x leverage—with just 1 yuan, you can build a 125-yuan position.

This is a level of leverage that almost no traditional market permits (stock futures top out at about 20-30x, Treasury bond futures at about 50-100x—but those require professional qualifications).

Feature Four: No Regulation at All / Very Weak Regulation

Many exchanges are registered on small Caribbean islands, in the Seychelles, in Dubai—with virtually no meaningful regulation. This means:

• The exchange may manipulate prices • The exchange may liquidate you in order to eat your margin • The exchange may simply abscond with the funds

10.3 The First Story: "Black Thursday," March 12, 2020

On March 11, 2020, the World Health Organization declared COVID-19 a "pandemic." Global stock markets began to crash.

The next day—March 12—the crypto market met the worst day in its history.

That day BTC fell from 7,900 dollars to 3,800 dollars—down 52% in a single day.

ETH fell from 195 dollars to 90 dollars—down 54% in a single day.

But it was not the spot holders who suffered most. Those who suffered most were the longs on perpetual contracts—

That day, the total amount liquidated across global crypto exchanges was about 3 to 4 billion dollars. On BitMEX alone, 1.6 billion dollars were liquidated within a few hours. Many people—at 50x, 100x leverage—went to zero within minutes.

Let me explain this in the language of the Kelly formula:

Suppose you went long on BTC at 7,900 dollars using 50x leverage.

That means: your "notional exposure" = 50 × your bankroll.

The liquidation price = 7,900 × (1 − 1/50) = 7,900 × 0.98 = 7,742 dollars.

In other words, all BTC has to do is fall from 7,900 to 7,742—a drop of 2%—and you go to zero.

On March 12, BTC fell 30%+ within minutes—you were long since wiped out, with no "chance to rebound" at all.

What would the Kelly formula tell you?

For an asset with annualized volatility of 80%, full-Kelly typically suggests leverage in the range of 2-5x.

50x leverage? That's not Kelly, that's suicide.

10.4 The Second Story: The May 19, 2021 Liquidation

In April-May 2021, BTC set what was then an all-time high of 64,000 dollars. The market was in a frenzy—everyone believed "this time is different," that Bitcoin would rise forever.

May 19—an ordinary Wednesday—BTC fell in a single day from 40,000 to 30,000, then rebounded to 35,000. At the trough, the drop exceeded 30%.

The total amount liquidated across the global crypto market that day: 9 billion dollars—the highest single-day liquidation record in crypto history.

Many retail investors who "added leverage to buy the dip" on May 11-18 were all wiped out on the 19th.

More frightening still was a fact discovered later: many of the "price flash crashes" on the large exchanges were manipulated by the exchanges themselves—in order to liquidate high-leverage positions and swallow retail investors' margin.

For example, one exchange was found to have deliberately engineered a brief "price anomaly" that day—triggering liquidations, then immediately restoring the price. This kind of "price manipulation" is an open secret in crypto markets.

10.5 The Third Story: May 2022—The Death of LUNA

If the crashes of 2020-2021 were merely "tragedies of leverage users," the LUNA collapse of May 2022 harmed everyone. What was LUNA?

It was the "governance token" of a crypto project called Terra Luna—and paired with it was a "stablecoin" called UST, designed to be worth 1 dollar forever.

UST's "stability" mechanism was very clever (and very fragile): when the price of UST was below 1 dollar, you could "burn" 1 UST for 1 dollar's worth of LUNA; when the price of UST was above 1 dollar, you could burn 1 dollar's worth of LUNA for 1 UST.

In theory this mechanism kept UST worth 1 dollar forever—because arbitrageurs would automatically balance it.

And LUNA promised UST holders 20% annualized interest—yes, 20% a year. This attracted millions of retail investors around the world—many people went all-in on UST with their houses, their pensions, their life savings—"it's a stablecoin anyway, always worth 1 dollar!"

LUNA's market capitalization reached 40 billion dollars in early 2022—one of the top 5 crypto projects by market cap at the time.

Then, beginning on May 9, 2022, came the crash—

• May 9: UST loses its peg for the first time (falling from 1 dollar to 0.99) • May 10: UST falls to 0.7 dollars—LUNA falls from 80 dollars to 30 dollars • May 11: UST falls to 0.3 dollars—LUNA falls to 5 dollars • May 12: UST falls to 0.1 dollars—LUNA falls to 0.5 dollars • May 13: LUNA falls to 0.0001 dollars—that is, from 80 dollars it lost 99.99988%

40 billion dollars of market capitalization—evaporated within 72 hours.

Millions of retail investors went completely to zero—many took their own lives. Reddit's r/Terra forum in those days was full of suicide-help posts. In South Korea (Terra's founder Do Kwon is Korean, and Korean retail investors held the most UST/LUNA), there were several publicly reported cases of suicide.

This was not "you only go to zero if you use 50x leverage"—here, spot holders who used no leverage at all also went to zero.

10.6 How the Kelly Formula Understands LUNA

Let me look at LUNA again through the Kelly formula.

If you invested in LUNA in early 2022—buying in at 80 dollars—what would the Kelly formula tell you?

The Kelly formula has one core premise: that you can objectively estimate your win probability and your odds.

But for an asset like LUNA: • What is the probability of a "stablecoin losing its peg"? Nobody knows—nothing on a comparable scale had ever happened before • Once the peg breaks, how much can you get out? Nobody knows—the market could go to zero in an instant • What is this asset's "intrinsic value"? Nobody knows—it has no cash flow, no assets, it is purely a product of "market confidence"

When all the key parameters are "nobody knows," what does the Kelly formula tell you?

The answer is: position = 0.

The Kelly formula has an implicit principle: when your estimates of the parameters are highly uncertain, the optimal position approaches zero.

But the vast majority of LUNA investors had no such awareness—they saw "20% annualized interest" and went all-in.

This is the deepest warning of the Kelly formula—betting in a market you do not understand at all is equivalent to not betting.

10.7 The Fourth Story: The Collapse of FTX, November 2022

On November 8, 2022—six months after the LUNA crash—FTX, the world's second-largest cryptocurrency exchange, was found to be insolvent.

FTX's founder, Sam Bankman-Fried (SBF)—a 30-year-old who had graduated in physics from MIT—was accused of misappropriating customer funds: he lent billions of dollars of customer funds to his own hedge fund, Alameda Research, to make high-risk crypto speculations—and Alameda had lost most of that money in the LUNA crash.

On November 8, the news broke. On November 9, FTX users began withdrawing en masse. On November 10, FTX suspended withdrawals—customers' money was frozen. On November 11, FTX filed for bankruptcy—customers completely lost the right to access their own funds.

At the time, FTX managed roughly 10 to 15 billion dollars in customer assets. Over 90% of that money went to zero in an instant—part of it was later recovered through bankruptcy proceedings, but many customers still have not gotten back all of their principal.

In November 2023, SBF was convicted on multiple counts including fraud, money laundering, and violating campaign finance law, and in March 2024 he was sentenced to 25 years in prison.

This is another profound warning from the Kelly formula—

Exchange counterparty risk—even though your assets are "in your account," the exchange can lock them up, misappropriate them, or zero them out at any time.

In the face of this risk, the Kelly formula offers no protection whatsoever. The only protection is not to hold that much.

10.8 The "Iron Laws of Kelly" for Crypto Markets

After the several great crashes of 2020-2022, professional crypto quant funds distilled a set of "iron laws of crypto Kelly":

Iron Law 1: Perpetual-contract leverage no higher than 5-7x

Reason: BTC's daily volatility is about 3-5%, and the liquidation distance should be at least 5 times the daily volatility (25%)—working backward, leverage ≤ 4x. For ETH or other altcoins, leverage should be even lower (2-3x).

Iron Law 2: Size positions with a "pessimistic estimate"

For every estimate of μ (expected return) and σ (volatility), make at least three—"optimistic, neutral, pessimistic"—and execute according to the "pessimistic" column. This absorbs estimation error.

Iron Law 3: No single coin above 20% of total capital

Even if you are most bullish on BTC—do not exceed 20%. Reason: a black swan (a regulatory crackdown, a technical bug, an exchange hack) could take any single coin to zero.

Iron Law 4: No single exchange above 30% of total capital

Reason: the FTX lesson—an exchange can fail. Spread your assets across 3-5 exchanges (ideally including a regulated U.S.-based exchange such as Coinbase).

Iron Law 5: Hold most long-term positions in cold wallets

Keep on each exchange only the funds you need for "active trading." The vast majority of long-term holdings should sit in a cold wallet whose private keys you control yourself—that way, no matter what happens to the exchange, your assets are safe.

Iron Law 6: Use the funding rate contrarily

A perpetual contract's funding rate reflects market sentiment. When the funding rate is extremely high (30%+ annualized), the cost of going long is extremely high—this is often a market top. When the funding rate is extremely low or negative, the cost of going short is extremely high—this is often a market bottom.

Iron Law 7: Always keep 30% in cash (USDT or USDC)

Reason: at the top of a bull market you need cash to take profits; at the bottom of a bear market you need cash to buy the dip—without cash, you can neither realize gains nor seize opportunities.

10.9 A Set of Numbers That Quiet You

Finally, a set of numbers.

The wealth cumulatively evaporated from the crypto market between 2017 and 2025:

• Cumulative drawdown across BTC's successive "bear markets": about 5-7 trillion dollars • Total market cap of the thousands of altcoins that went to zero: about 1-2 trillion dollars • Customer losses caused by failed exchanges (Mt. Gox, Quadriga, FTX, and others): about 20 billion dollars • Cumulative liquidations of perpetual contracts: estimated at more than 1 trillion dollars

Most of this money flowed to three kinds of people:

• Early coin hoarders: bought BTC below 1,000 dollars and sold above 50,000—some early individual investors got rich • Quant funds: using a professional Kelly fraction + strict risk control, they profited steadily over the long run • Exchanges: living off fees and the margin of liquidated users —they are the crypto market's most stable winners

And the vast majority of ordinary retail investors—with no edge, no discipline, no Kelly awareness—became the conveyors of that wealth.

10.10 The True Theme of This Chapter

Writing this far, I want to pause.

The 9 chapters before this all told of the "winners" of the Kelly formula. This chapter tells only of losers—with no names, no stories, only numbers.

Why?

Because the true power of the Kelly formula lies not in how much it can make you win, but in which losses it can help you avoid.

Of the 40 billion dollars lost in the LUNA crash, the overwhelming majority could have been avoided—if those retail investors had:

• Known that a "stablecoin losing its peg" is a real risk • Not gone all-in on a single asset with all their capital • Set aside 30%+ in cash to weather the unexpected • Not used leverage • Not placed large bets on assets they did not understand

These are all basic principles of the Kelly formula. But they did not know them. Or, knowing them, they chose to ignore them.

Every lesson the Kelly formula taught over seventy years—Thorp at blackjack, PNP on Wall Street, Bill Gross at PIMCO, Buffett at Berkshire, Benter at the Hong Kong races, Walters and Voulgaris in sports betting—those lessons replay again and again in the crypto market.

At every crash, someone says: "This time is different." After every crash, new retail investors arrive and repeat the same mistakes.

And that 1956 formula—still waiting quietly inside a Bell Labs paper—waits for the next person willing to listen to what it has to say.

Chapter Summary: The Crypto Market's Most Profound Confirmation of the Kelly Formula

First: the true value of the Kelly formula lies in "preventing ruin"—the crypto market, with 5 trillion dollars evaporated, has proven the iron law that "overbetting = inevitable ruin."

Second: leverage is a slow poison—50x leverage on an asset with 80% volatility is expected to go to zero within a few days. The Kelly formula suggests a maximum crypto leverage of 2-5x.

Third: exchange counterparty risk—the FTX lesson: your assets may go to zero because an exchange fails, and the Kelly formula cannot defend against this risk—the only defense is to spread across multiple exchanges + cold wallets.

Fourth: for an asset you do not understand at all, the optimal position = 0—the LUNA lesson: when all the key parameters are "nobody knows," the Kelly formula's solution is not to bet.

Part Two, "Into the Field," ends here. We have looked at 7 domains and more than a dozen protagonists—people who met this 1956 formula, worked with it, sometimes defied it, and were sometimes saved by it.

In the next part, "Controversy and Reflection," we will step into another conversation—the economics profession's opposition to the Kelly formula. The loudest voice of this opposition came from a Nobel laureate in economics, who wrote an entire paper in monosyllabic English words to humiliate the supporters of the Kelly formula.

His name was Paul A. Samuelson.

Part Three · Controversy and Reflection

Chapter 11 · The Nobel Laureate's Monosyllabic Paper

11.1 A Paper Written in a Single Syllable

1979. Cambridge, Massachusetts. The Department of Economics at MIT.

Paul Anthony Samuelson—the "pope" of American economics, winner of the 1970 Nobel Prize in Economics, sixty-four years old—laid a typescript on his secretary's desk:

"Send this off for me to the Journal of Banking & Finance."

His secretary picked it up and glanced at the title—

"Why we should not make mean log of wealth big though years to act are long."

She paused.

The entire title—apart from the prepositions and conjunctions—every word had only one syllable.

She turned to the body of the text.

The whole paper was written in a single syllable.

Every sentence, every paragraph, every word—all the simplest one-syllable English words. Except for the signature at the end.

She looked up at Samuelson, her eyes full of bewilderment.

Samuelson said with a smile: "This is for those Kelly supporters who 'can't follow the hard math.'"

This was the most dramatic moment in the seventy-year history of the Kelly formula debate—the highest authority in the field of economics, using the most childish English, humiliating the supporters of the Kelly formula with the utmost elegance.

11.2 Who Was Samuelson

To understand why Samuelson's attack was so lethal, one must first understand who he was. Paul Samuelson was born in 1915 in Gary, Indiana, into a family of Russian Jewish immigrants. He entered the University of Chicago at sixteen, earned his bachelor's degree at twenty, completed his master's at MIT at twenty-one, and received his doctorate at twenty-six.

His doctoral dissertation, Foundations of Economic Analysis (published in 1947), is considered one of the most important works of twentieth-century economics—it rewrote the whole of economics using mathematics (in particular, calculus and optimization theory).

He was the first American to win the Nobel Prize in Economics (1970).

The undergraduate textbook he wrote, Economics (first edition 1948), sold more than four million copies in the twentieth century—the best-selling economics textbook in history.

He taught in MIT's economics department for more than sixty years. Most of the modern methods of analysis in economics—including the rigorous mathematical definition of "utility function," "welfare," and "consumption choice"—were laid down by him.

In short: he was the Newton of economics.

11.3 Samuelson's Three Objections to the Kelly Formula (1969, 1971, 1979)

Samuelson had opposed the Kelly formula fiercely beginning in the 1960s. His opposition was not a one-time affair, but ran through ten years—three principal papers:

The first time: 1969, Review of Economics and Statistics

In this paper Samuelson set out his central argument for the first time:

"Maximizing log(W) over the long run—that is, the Kelly strategy—is not what a rational investor truly wants to do.

It is merely one particular solution for the logarithmic utility function (log utility)—and there is no reason whatsoever why the logarithmic utility function should be superior to any other utility function.

If you prefer a different degree of risk aversion, you should use a different strategy—rather than worshipping Kelly."

This was a purely theoretical attack: Samuelson did not deny the mathematical correctness of the Kelly formula, but he attacked its uniqueness—it is not the one and only optimum.

The second time: 1971, PNAS

In 1971, Samuelson published a paper in the Proceedings of the National Academy of Sciences (note—the very same journal in which Thorp had published his blackjack paper years earlier), with a sharper title: "The 'fallacy' of maximizing the geometric mean in long sequences of investing or gambling"

The word "fallacy" is a very hard word—it is the harshest accusation an economist can make.

In this paper Samuelson used six pages of rigorous proof to argue:

• The Kelly strategy maximizes the long-run growth rate—this is true • But "maximizing the long-run growth rate" does not equal "maximizing a rational investor's utility" • An investor who is extremely averse to extreme losses should use a strategy more conservative than Kelly • An investor who is risk-neutral (rare, but theoretically possible) should use a strategy more aggressive than Kelly • There is no a priori reason to say that "logarithmic utility is correct"

The third time: 1979, the monosyllabic paper

By 1979, Samuelson had lost his patience.

He had noticed—despite the fact that he and other economists had published objection after objection—that the supporters of the Kelly formula (particularly Thorp, MacLean, Ziemba, and others) were still publishing a great many works claiming that "Kelly is optimal."

He felt that these supporters could not follow complicated economics—so he decided to state his argument once more, using the simplest English.

And so came that monosyllabic paper—

"Why we should not make mean log of wealth big though years to act are long"

Translated into Chinese it reads:

"为什么我们不应当把财富的对数均值做大，即使我们要做很多年"

This is how the paper opens—you can feel this sixty-four-year-old Nobel laureate grinning wickedly as he types at his typewriter:

"He who acts in N plays to make his mean log of wealth as big as it can be made will, with chance one as N grows, beat me who acts to meet my own tastes."

(If in N bets you pursue the largest mean log of wealth, then so long as N is large enough, you will almost surely beat me—but I am acting according to my own tastes.) "Yet, when N grows big, my mean log score may stay far behind his. Why? My tastes are not your tastes. I write large in odes the worth I place on each step's win or loss. To bet so as to make my long log score big is to bet as one whose log of worth is what he aims at. That is not my aim."

(True, when N is large my mean log of wealth will fall far behind yours—but that is because our tastes differ. I assign greater weight to each step's win or loss—to bet according to "my tastes" is not to maximize the log of wealth. That is not my aim.)

Note—every word is a single syllable: he, who, acts, plays, make, mean, log, wealth, big, can, be, made, will, with, chance, one, grows, beat……

11.4 What Samuelson Was Attacking

Let us restate Samuelson's argument in Chinese:

Point 1: The Kelly formula assumes that the investor maximizes logarithmic utility log(W). But logarithmic utility is not universal—it is only one among a great many utility functions.

What is a utility function? Economists use it to represent "your preferences over wealth":

• If your utility function is $u(W) = W$, you are risk-neutral—you care only about expected wealth and are indifferent to volatility • If your utility function is $u(W) = \log(W)$, you are a Kelly investor—you care about the logarithmic expectation of wealth • If your utility function is $u(W) = -1/W$, you are highly risk-averse—the slightest loss causes you great pain

The Kelly formula is optimal only in the case $u(W) = \log(W)$. For other utility functions, the optimal solution is the CRRA formula:

$\mu - r$

$$f^* =$$

$\gamma\sigma^2$

where γ is your "coefficient of risk aversion." When $\gamma = 1$, this is the Kelly formula (logarithmic utility). When $\gamma = 2$ it is "half Kelly," and when $\gamma = 4$ it is "quarter Kelly."

Point 2: Maximizing the long-run growth rate does not equal rational maximization.

The core of Samuelson's objection is this: even if the Kelly strategy almost surely wins in the long run (Breiman 1961, Property 2), this still does not mean that a rational investor ought to choose it.

Why? Because "almost surely wins" ignores "how badly you lose when you lose." The Kelly strategy will, much of the time, experience enormous drawdowns (with full Kelly, the probability of a 50% drawdown is about 33%). An investor extremely averse to drawdowns—might well prefer to win a little less, but with far smaller drawdowns. This is precisely the fundamental reason for using CRRA with $\gamma > 1$ (half Kelly or lower).

Point 3: Kelly supporters often misuse the concept of "the long run."

Another point in Samuelson's attack—a very clever one—concerns "the long run."

Kelly's "long-run optimality" requires $N \to \infty$ (the number of bets tending to infinity). But in reality:

• Anyone's investing time is finite (at most 60–80 years) • Economic mechanisms change (monetary systems, policies, technological revolutions) —the so-called "infinite repetition of the same game" does not hold in reality

Samuelson's counter-question is: if "the long run" is 50 years—rather than infinite years—then is the Kelly strategy really optimal?

The answer is: not necessarily. Within a 50-year window, the effect of the degree of risk aversion is large—it may make half Kelly, or even quarter Kelly, more optimal.

11.5 Thorp's Counterattack

The moment Samuelson's monosyllabic paper appeared in 1979, Thorp was infuriated.

By that time Thorp was already forty-seven, a core manager of PNP, having made tens of millions himself. He was not an economist, but he was a practitioner of the Kelly formula—he knew that the Kelly strategy really did work in his own account.

He wrote several responses—the most important of which was his 1984 paper, The Kelly Criterion in Blackjack, Sports Betting, and the Stock Market.

The key points of Thorp's rebuttal:

First: Samuelson is right—Kelly is not universally optimal. But most practical investors—especially those not skilled at precisely articulating their own risk preferences—default to something close to logarithmic utility.

"Ask an ordinary person: would you accept a wager with a 50% chance of doubling and a 50% chance of halving?

Most people would say: no—it feels as though the risk outweighs the reward.

That is precisely the response of log utility—log(2W) + log(0.5W) = log(W^2) - log(2W) = -log(2) < 0. Logarithmic utility rejects this wager.

This means: for most people, logarithmic utility is a reasonable approximation. And the corresponding strategy is Kelly." Second: The "long run" problem Samuelson attacked can be solved through "fractional Kelly."

"Samuelson is right—full Kelly's drawdowns are too large.

But the supporters of the Kelly formula have long been using fractional Kelly—half Kelly, quarter Kelly, one-eighth Kelly.

This is the concrete form of the 'more drawdown-averse' solution that Samuelson wanted.

Samuelson is attacking a straw man—he is attacking the extreme position that 'one must use full Kelly,' whereas real Kelly users never do such a thing."

Third: Samuelson himself is actually using Kelly.

"Samuelson's own portfolio in his later years consisted mainly of diversified stocks—which is essentially the solution for CRRA γ ≈ 2–4, equivalent to quarter to half Kelly.

What he opposes is 'the extreme form of Kelly,' but what he actually uses is a discounted version of Kelly.

This is exactly what I and PNP are doing."

This debate later became a celebrated case in the history of economics. Each side held to its own view—neither fully persuaded the other, but both revised their extreme positions.

11.6 The Present Consensus: The CRRA Spectrum

After more than thirty years of debate, the fields of economics + quantitative finance arrived at a compromise consensus:

The Kelly formula is not universally optimal, but it is the particular solution at γ=1 within the CRRA utility spectrum.

In practice one should choose γ according to one's own degree of risk aversion, and the corresponding optimal strategy is fractional Kelly:

γ Name Kelly fraction Suited investor γ = 0.5 Risk-seeking 2 × Kelly almost no one γ=1 Log utility full Kelly the theoretical investor γ=2 Moderate risk aversion half Kelly most hedge funds γ=4 High risk aversion quarter Kelly most retail investors, pension funds γ = 10 Extreme risk aversion 1/10 Kelly conservative investors, near retirement

The most common in practice is γ = 2–4—that is, "half Kelly to quarter Kelly."

Samuelson is right—pure Kelly is not suited to most people. Thorp and MacLean are also right—fractional Kelly is suited to most people. The debate between the two men caused the Kelly formula to evolve from "full Kelly" into a "fractional Kelly spectrum"—and this is the true contribution of the debate.

11.7 Samuelson's Later "Reconciliation"

In his later years—he lived to ninety-four, dying in 2009—Samuelson's attitude toward the Kelly formula gradually softened.

In interviews and papers in the 2000s he acknowledged on several occasions:

"The Kelly formula is reasonable in certain specific settings— especially: (1) when the investor truly has a logarithmic utility preference (2) when the investor intends to compound over the long run (10 years or more) (3) when the investor can withstand a 30–50% drawdown

But as a universally optimal strategy, Kelly is still not correct—it is only the solution at the particular point CRRA γ=1."

Samuelson also acknowledged that the fractional-Kelly theory of Thorp, Markowitz, and others was a reasonable response to Samuelson's objection—fractional Kelly is in fact a compromise shared among the three of them: Samuelson–Markowitz–Thorp.

When Samuelson died in 2009, MIT's economics department published an obituary—which made almost no mention of his debate with Kelly. Heated though the debate was, it was only a very small part of Samuelson's overall academic contribution.

After Samuelson's death, Thorp wrote a eulogy. He said:

"Samuelson and I argued about the Kelly formula for forty years.

Now he has passed away.

I think we were each half right.

He was right that: the Kelly formula is not a universal truth.

I was right that: under a reasonable discount, the Kelly formula is the best guide for most practical investors.

Our debate made this formula more useful—and made us both more humble.

This is the best possible ending for an academic debate."

11.8 The Lessons of This Chapter

Writing this far, I want to pause—

What lessons do Samuelson and Thorp's debate give us?

The first lesson: the Kelly formula is not a law of physics. It is a mathematical model with specific assumptions—that your preference is logarithmic utility, that you can estimate the parameters precisely, that you can repeat over the long run. These assumptions are not all true—and so the Kelly formula cannot be applied blindly.

The second lesson: economists and practitioners can offer complementary insights from different angles.

What Samuelson offered is theoretical sobriety—reminding us of the boundaries of Kelly. What Thorp offered is practical wisdom—reminding us of the power of Kelly.

The two together—fractional Kelly + CRRA choice—are the tools used by professional investors today.

The third lesson: academic debate makes truth more complex, and also more accurate.

Had it not been for Samuelson's attack, the Kelly formula might still have been stuck at "full Kelly"—and the LTCM 30-times-leverage collapse of 1998, which we are about to examine, might have been "full Kelly in real life."

It was precisely Samuelson's attack that forced the Kelly school toward "fractional Kelly"—and over the past forty years fractional Kelly has saved the lives of countless investors.

Academic opponents are often future collaborators in being proven right.

Chapter Summary: Samuelson's Three Gifts to the Kelly Formula

First: he questioned the extreme position that "Kelly is universally optimal"—forcing the Kelly school to admit that it is only the particular solution at CRRA $\gamma=1$.

Second: he redefined the position of the Kelly formula in the language of the CRRA spectrum—giving "fractional Kelly" a theoretical foundation.

Third: through the vehicle of a monosyllabic paper he set a paradigm for the field of economics—that academic debate can be serious, can be humorous, and can be effective.

In the next chapter, we will look at a story of not heeding Samuelson's warning—the hedge fund LTCM, founded by two Nobel laureates in economics (one of them the Scholes of the Black-Scholes formula), which used 30-times leverage and, within three months in 1998, evaporated $4.6 billion, nearly triggering the collapse of the global financial system.

The story of LTCM tells us—no matter how high your IQ, no matter how many Nobel Prizes you have won, the price of violating the Kelly formula is always the same.

Chapter 12 · The Nobel Halo at Thirty-to-One Leverage

12.1 An Opening That Astonished Everyone

February 1994. Greenwich, Connecticut.

A new hedge fund announced its founding — Long-Term Capital Management (LTCM for short).

Its roster of partners made all of Wall Street hold its breath:

• John Meriwether — Wall Street's legendary bond trader, Salomon Brothers' "King of Bond Arbitrage" in the 1980s, former head of Salomon's fixed-income division • Myron Scholes — the Scholes of the Black-Scholes option-pricing formula, professor at Stanford's business school • Robert C. Merton — founder of dynamic hedging theory, professor at Harvard Business School • David Mullins — former Vice Chairman of the Federal Reserve • Eric Rosenfeld — former chief arbitrage strategist at Salomon Brothers • William Krasker — former professor at Harvard Business School

— plus several other of the finest bond-arbitrage experts on all of Wall Street.

This was the most luxurious hedge-fund team ever assembled.

Myron Scholes and Robert Merton jointly won the Nobel Prize in Economics in 1997 — the third year after LTCM's founding (Fischer Black had already died in 1995, and under the rules the prize cannot be awarded posthumously).

LTCM raised $1.3 billion — at the time the largest hedge-fund launch in history.

It promised:

• Net returns of 20-30% per year • Volatility no more than half that of the S&P 500 • Never a loss exceeding 20%

Over the next four years, LTCM delivered on nearly all its promises:

• 1994: net return 28% • 1995: net return 43% • 1996: net return 41% • 1997: net return 17% (the year Scholes and Merton took the Nobel Prize)

The four-year average annualized net return was about 32% — more than twice the S&P 500 over the same period. LTCM became the star of the global hedge-fund industry — investors lined up to get in, banks scrambled to lend to it, and the news media hailed it as "the revolution of finance."

Then came 1998.

12.2 LTCM's Strategy: Relative-Value Arbitrage

What LTCM did was, in essence, the same as PNP — seek out mispricings and lock in risk-free profit through hedged portfolios.

Its core strategy was called "relative value arbitrage." Consider a few examples:

Example 1: The "on-the-run vs. off-the-run" spread in U.S. Treasuries

The U.S. government issues new 30-year Treasury bonds every year. The most recently issued one is called "on-the-run" — extremely liquid, and used by all traders for hedging. The one that has been outstanding for a few months is called "off-the-run" — slightly less liquid, but essentially identical to the on-the-run (both are U.S.-government-backed, 30-year).

In theory, the two prices should be exactly equal. In practice, the on-the-run, because of a "liquidity premium," is a few basis points more expensive than the off-the-run (say, 5-10 bps).

LTCM's move: go long the cheap off-the-run and short the expensive on-the-run. When, a few months later, the new on-the-run comes out and the old on-the-run becomes an off-the-run, the prices naturally converge — and LTCM locks in 5-10 bps of profit.

The profit on a single trade is tiny (5-10 bps). But through massive repetition — LTCM could do hundreds of similar trades — the total profit is considerable.

Example 2: The spread between the government bonds of different European countries

In the late 1990s, Europe was preparing for the unified currency, the euro. The German mark, French franc, Italian lira, Spanish peseta, and others were gradually converging.

LTCM's judgment: Italian bond yields would converge toward German bond yields (because, after currency unification, interest rates would level out). So it went long Italian bonds and short German bonds — locking in the "convergence profit."

Example 3: Convertible-bond arbitrage

The same strategy as PNP — go long the undervalued convertible bond and short the corresponding stock.

12.3 LTCM's Leverage Problem

All of LTCM's strategies shared one common trait: the profit on any single trade was minuscule (a few basis points).

So how could the fund earn 30% annualized? With leverage.

How much leverage did LTCM use?

The official figures at the start of 1998:

• Equity: $5 billion • Total positions on the balance sheet: $125 billion • Notional value of off-balance-sheet derivatives: about $1.25 trillion

On-balance-sheet leverage: 25 times. Total notional leverage including derivatives: 250 times.

This is a staggering number — 5 to 50 times that of a typical hedge fund of the era (5-10 times).

Why did LTCM dare to use such high leverage?

LTCM's logic ran like this:

"The risk in each of our trades is extremely low — because what we do is hedged portfolios — the long and short positions are highly correlated in price, and the net exposure is small.

Even if each trade earns only 5-10 bps and the risk per trade is tiny, so long as we repeat it hundreds of times, the total profit is still considerable.

Our 'effective leverage' — that is, the leverage truly facing market risk — is far below 25 times — roughly equivalent to 3-5 times in traditional asset allocation."

This logic has a certain reasonableness. Under normal market conditions, LTCM's strategy really was extremely stable — the continuous profits of 1994-1997 proved as much.

But the logic of the Kelly formula would tell you:

This logic holds in a "normal market." But the market is not always normal.

In times of crisis, correlations shift abruptly — two positions that were supposed to hedge each other may suddenly move in the same direction. At that point your "hedged portfolio" is no longer a hedged portfolio; it becomes a doubled exposure.

25-times leverage + crisis correlation = catastrophe.

12.4 LTCM Did Not Use the Kelly Formula

This section is the most important part of this chapter.

Among LTCM's partners were Myron Scholes and Robert Merton — two Nobel laureates in economics — and Merton himself was precisely the mathematical economist who, in 1969-1971, extended the Kelly formula to continuous time!

We discussed in Chapters 2 and 3 how the Merton formula $\pi^* = (\mu - r) / \sigma^2$ is the continuous-time version of the Kelly formula.

In other words: among LTCM's partners was the inventor of the "continuous-time version of the Kelly formula."

So did LTCM use the Kelly formula?

The answer is: no. Or, more precisely — they used a "hacked-up version" of Kelly, but that hack seriously violated the core spirit of the Kelly formula.

Let us look at what LTCM actually used:

First: they used "VaR-based" risk control — based on Value-at-Risk — not on the Kelly formula.

VaR was the mainstream risk-control tool on 1990s Wall Street. It tells you "with 95% probability, how much you can lose at most tomorrow." LTCM used VaR to set a ceiling — for example, "VaR shall not exceed 5% of total capital."

But VaR has a fatal flaw: it assumes the market is normally distributed. A loss within the 95% probability range — that number tells you about the "normal case." It tells you nothing at all about "how large the loss is in the 5% tail case."

The Kelly formula (especially Merton's own 1971 version) carries an implicit warning: when fat-tail risk exists, one should sharply reduce the estimate of π^*. But LTCM ignored this warning.

Second: they used "historical variance" to estimate σ — and historical variance severely underestimated the true tail risk.

LTCM estimated σ from the past five years of market data. But 1994-1997 were among the calmest years in global financial markets — so the σ estimate was very low. Plugged into the Merton formula $\pi^* = (\mu - r) / \sigma^2$, π^* came out very large — meaning high leverage was permissible.

But the true σ — accounting for long-run history plus crisis scenarios — was far larger than that estimate.

One of the Kelly formula's core warnings is: using an "optimistic σ" leads to severe overbetting. LTCM applied this warning not at all.

Third: they assumed "normal correlations" — whereas correlations shift abruptly in a crisis. LTCM's hedged portfolios relied on stable correlations between assets. For example, they assumed the correlation between "U.S. Treasuries vs. Italian bonds" was about 0.7.

But after the Russian crisis of 1998 — the correlations of all assets suddenly spiked to 0.95-1.0. The long and short positions that were supposed to hedge each other lost money at the same time — something LTCM had not anticipated.

The multi-asset version of the Kelly formula (in Chapter 11 we discussed the Ledoit-Wolf shrinkage covariance) requires the use of a stress-tested Σ — that is, one must account for "crisis correlations." LTCM did not do this.

12.5 The Collapse of 1998

August 17, 1998. A Monday.

The Russian government announced: the ruble would be devalued, and its foreign debt defaulted.

This was a once-in-a-century event — a major industrial nation defaulting outright. Global capital markets fell into panic.

What followed was something LTCM had never considered in its models:

First: all "safe assets" were bought up — the yields on U.S. Treasuries and German bonds fell sharply.

Second: all "risky assets" were sold off — Italian bonds, emerging-market debt, convertible bonds, and high-yield bonds plunged.

Third: LTCM's hedged portfolios suddenly failed across the board —

• the "undervalued assets" they were long (Italian bonds, off-the-run Treasuries, etc.) all fell • the "overvalued assets" they were short (German bonds, on-the-run Treasuries, etc.) all rose

In other words, both sides lost money at once — rather than the "hedging on both sides" they had designed.

Add 25-to-30-times leverage — and the losses were mercilessly magnified.

August 1998: LTCM lost $1.9 billion in a single month (38% of total equity). September 1998: LTCM lost another $2.7 billion — a cumulative loss of $4.6 billion over two months — exceeding 90% of total equity.

By the end of September 1998, LTCM's remaining equity was only about $400 million — while its total notional positions still exceeded $1 trillion.

If LTCM went under — its hundreds of counterparty banks (including Merrill Lynch, Lehman, JPMorgan, UBS, and nearly every top Wall Street bank) would face enormous losses simultaneously. The global financial system might collapse.

12.6 The Federal Reserve's Emergency Rescue

On September 23, 1998, the Federal Reserve urgently convened 14 top Wall Street banks for an unprecedented meeting:

The meeting was held at the Federal Reserve Bank of New York. The substance of the meeting was: these 14 banks would jointly put up $3.65 billion to take over all of LTCM's positions, averting the shock to the market that a forced liquidation of LTCM would cause.

After 12 hours of negotiation, the plan was finally approved. LTCM was in effect taken over jointly by the Federal Reserve and Wall Street — the partners retained a small stake but lost almost all management control.

After the takeover, the market gradually calmed — LTCM's positions were liquidated in an orderly way, without triggering a larger crisis.

But LTCM's story was over.

The final tally:

• Personal losses of LTCM's partners: estimated at about $1.8 billion (all of the partners' prior profits wiped out, plus large personal-loan losses) • Losses of early investors: about $1.7 billion • Losses of the rescuing banks: about $400 million to $1 billion (partly recovered through liquidation) • Overall market impact: estimated in the hundreds of billions of dollars (including all knock-on effects)

The two Nobel laureates — Scholes and Merton — saw their personal net worth fall from hundreds of millions of dollars to tens of millions (still wealthy, but far below their peak). Scholes later opened a new fund, Platinum Grove, which lost 38% in 2008 — failing yet again.

12.7 The True Lesson of LTCM

LTCM's story has been written up in countless books — Roger Lowenstein's When Genius Failed is the most classic of them. But most books frame it as a story of "even Nobel laureates can fail" — and that is only the surface lesson.

The deeper lesson runs like this:

First: Merton himself invented the continuous-time version of the Kelly formula, but LTCM did not use it.

This is a bewildering fact. Merton's 1971 paper tells you clearly that $\pi^* = (\mu - r) / \sigma^2$ — yet LTCM's actual leverage far exceeded the maximum that formula recommends (even using the most optimistic parameter estimates).

Why did Merton himself betray his own formula? The answer is probably this: he believed his own model could "surpass the conservatism of the Kelly formula" — through precise hedging, precise covariance estimation, and precise tail-risk modeling — he believed he could safely use 30-times leverage.

This is the mistake economists most often make — trusting their own model to be more precise than the Kelly formula.

Second: LTCM completely ignored the "fractional Kelly" tradition of the Kelly formula.

We have seen earlier — Thorp, Bill Gross, Buffett — every truly successful user of Kelly used "fractional Kelly" — typically 1/4 to 1/2 Kelly. This is to absorb estimation error.

LTCM used "super-Kelly" — even if their parameter estimates had been accurate, their leverage exceeded the maximum recommended by the Kelly formula. Any error in parameter estimation was mercilessly magnified by geometric compounding.

Third: LTCM did not account for "crisis correlations."

The multi-asset version of the Kelly formula requires a stress-tested Σ. LTCM used "normal-period correlations" — leaving no margin at all for the possibility that "in a crisis all correlations go to 1."

12.8 If the Kelly Formula Had Been Obeyed, What Would Have Become of LTCM?

Let us engage in a counterfactual —

Suppose that when LTCM was founded in 1994 it had obeyed the fractional-Kelly principle of the Kelly formula — that is, used 1/4 to 1/2 Kelly:

• Its leverage would have dropped from 25 times to 5-10 times • Its annualized net return would have dropped from 30% to 8-15% • Its maximum drawdown in the 1998 crisis would have dropped from over 90% to 20-30% — and it would still have survived

It would still have been profitable. It would still have been regarded as a top hedge fund. But it would not have gone bankrupt. Merton and Scholes would still have had personal net worth in the hundreds of millions — rather than tens of millions.

More importantly — the global financial crisis of 1998 would have been greatly eased — LTCM would not have needed a Federal Reserve rescue, and the market would not have fallen into panic.

All of this — merely by cutting leverage from 25 times to 5 times.

This is exactly what Samuelson warned of in his 1979 monosyllabic paper —

"Do not chase the maximal mean of log wealth when it is not necessary."

Samuelson was right. LTCM's tragedy is the real-world version of Samuelson's warning.

12.9 A Comparison Worth Pondering

Let me use a table to compare PNP and LTCM — both were hedge funds applying the Kelly formula, so why were their fates so different:

Princeton-Newport Partners Long-Term Capital Management Year founded 1969 1994 Founders Edward Thorp Meriwether + Scholes + Merton (mathematician) (Nobel laureates) Strategy Convertible arb + Relative-value arbitrage statistical arb Kelly usage 1/4 - 1/2 Kelly Super-Kelly (25-times leverage) (conservative) Single-position 2% of total capital 5-10% of total capital (far above this ceiling once leverage is included) Leverage about 5-8 times about 25-30 times Covariance stress-tested Σ normal-period Σ estimation 1987 Black roughly flat (not yet founded) Monday that month 1998 Russian profit of about loss of 92%, rescued crisis 12% that year Final outcome 19 years with no collapsed within 5 years, losing quarter, taken over by the Fed shut down voluntarily

The difference between the two funds — is not "who was smarter" — it is that Thorp heeded the Kelly formula and Meriwether did not.

12.10 What This Chapter Really Means

LTCM's story is not "even Nobel laureates can fail" — that is only the surface.

LTCM's story is: no IQ can substitute for the discipline of the Kelly formula.

You can be a doctor of physics, you can be a Nobel laureate, you can invent the Black-Scholes formula, you can hold every academic honor there is to hold —

If you violate the core discipline of the Kelly formula: • using full-Kelly or super-Kelly • using optimistic parameter estimates • ignoring crisis correlations • single positions too large

you will, at some point, be punished by the market.

This is not about "luck." It is the mathematical inevitability of geometric compounding.

LTCM's 30% return every year from 1994 to 1997 was a mathematically reasonable result — in that four-year "normal market," their strategy really did have an edge.

LTCM's 90% loss in three months in 1998 was likewise a mathematical inevitability — 25-times leverage + an abrupt shift in crisis correlations = no matter how smart you are, the result is the same.

This is the most profound lesson the Kelly formula has given the economics profession.

Chapter Summary: LTCM's Four Warnings to Posterity

First: leverage is the "amplifier" of the Kelly formula — it amplifies profits, and it amplifies losses. Any leverage above what the Kelly formula recommends is slow suicide.

Second: normal-period parameters ≠ crisis-period parameters — LTCM used "calm-period" parameters from 1994-1997 to estimate risk, and the result failed across the board in the "crisis period" of 1998.

Third: hedged portfolios "come unlocked" in a crisis — all correlations go to 1, the hedge vanishes, and the exposure doubles.

Fourth: IQ cannot substitute for discipline — two Nobel laureates + Wall Street's strongest team + a perfect mathematical model still could not redeem the cost of "violating the discipline of the Kelly formula."

The next chapter — the last chapter of this book — will bring us back to the present.

Today, in 2026, AI large language models (such as Claude, GPT, Gemini) can already help investors make decisions in real time.

In the age of AI, how should the Kelly formula be used?

The answer is not "let the AI make decisions for you" — but to let the AI be your "adversary."

Part Four · The Modern Era

Chapter 13 · Asking an AI to Write a Letter of Dissent

13.1 A Morning in 2026

A Tuesday morning in March 2026. A retail investor in Shanghai, forty-three years old—let us call him Old Wang.

He turns on his computer and stares at the BTC chart—overnight, BTC has suddenly jumped 8%, from $82,000 to $88,500. His fingers are itching—he wants to add to his position.

He runs the numbers on his account: 1 million RMB, of which 300,000 is already allocated to crypto (about 30%).

He pauses in front of the screen for a moment, then opens Claude—over these past two years, using Claude to help him make trading decisions has become a habit.

He types:

"Claude, I'm about to increase my BTC position from 300,000 to 500,000—which is to say, raise BTC from 30% of my portfolio to 50%. Please analyze this decision using the Kelly formula."

Claude replied—but the way it replied was not what he expected—

"I can help you compute the Kelly formula. But before computing it, I want to ask you a few questions first—

1. What do you estimate BTC's expected return over the next 12 months to be? Give me a specific number. 2. What do you estimate BTC's volatility (annualized) over the next 12 months to be? 3. What is the basis for your estimate of the win rate? 4. How much maximum drawdown can you tolerate? 5. If BTC falls 40% within a week after you add to your position, what will you do?

Please answer each one. Until you answer clearly, I cannot give you a Kelly-formula recommendation—because 'plugging optimistic parameters into Kelly' is the reason most people go broke."

Old Wang was stunned.

He had assumed Claude would immediately hand him a number—"Kelly recommends you bet 40%" or "Kelly recommends you bet 20%"—a simple, executable answer.

But Claude gave no number. Claude forced him to first answer questions he had almost never seriously considered.

He tried to answer: "1. BTC expected return, say 50%—it's been this strong lately. 2. Volatility? I don't know, 60% maybe. 3. Win rate? I feel very confident, 70%. 4. Maximum drawdown… I should be able to tolerate 30%? 5. A 40% drop… I'd cut my losses, I guess."

Claude replied:

"Thank you. Now I can give you three analyses—

Analysis 1: Your parameters have a serious overconfidence bias

You estimate BTC's expected return over the next 12 months at 50%. That is an extremely optimistic estimate—the median of BTC's historical 12-month rolling returns is about 10–15% (including extreme bull-market years). Your estimate is 3–5 times higher than the historical median.

You estimate the win rate at 70%. This is a common expression of overconfidence among BTC investors—the true 'probability of making money' (the frequency of positive 12-month rolling returns) is about 55–60%. You have overestimated it by 10–15 percentage points.

Analysis 2: Plugging in your parameters, Kelly recommends a position of nearly 100%

With $\mu = 50\%$, $\sigma = 60\%$, $r = 2\%$, the continuous-time Kelly formula gives $\pi^* = (0.50 - 0.02) / 0.36 = 133\%$—which is to say, 'by your optimistic estimates,' Kelly recommends you borrow money to add to your BTC.

But this is obviously an overestimate.

Analysis 3: Recomputing with 'pessimistic estimates'

Let me use more realistic parameters— - $\mu = 15\%$ (the historical median) - $\sigma = 80\%$ (crypto's true volatility, tails included) - $r = 2\%$ - $\gamma = 4$ (you said 'a 30% drawdown is intolerable,' which corresponds to CRRA $\gamma = 4$)

$$\pi^* = (0.15 - 0.02) / (4 \times 0.64) = 5\%$$

Which is to say, quarter Kelly recommends a BTC position of 5%—not 30% or 50%.

Your current 30% position is already 6 times the quarter-Kelly recommendation. Adding to 50%—would bring your actual position to 10 times the quarter-Kelly recommendation.

Analysis 4: The letter of dissent

If you increase your position to 50% and BTC falls 40% (an event that has happened 5 times within a 12-month window in BTC's history), your loss would be: 50% × 40% = 20% of total assets.

You said you can tolerate a 30% drawdown—this single trade would use up two-thirds of your 'drawdown budget.' Any other adverse event on top of it—and you break through the boundary of what you can bear.

You said, 'If it falls 40%, I'll cut my losses.' The problem is: you won't cut before it falls 40%—you will keep telling yourself 'wait a little longer' through the fall of 10%, 20%, 30%, and finally cut your losses at the very bottom of the 40% drop—perfectly realizing the maximum loss.

This is the core of everything the Kelly formula warns against 'overconfidence.'

My recommendation:

• Don't add to your position. Keep your existing 30% position—this is already the limit of your 'risk budget.' • If you truly want to add, set the total ceiling for the position at 35–40%—do not exceed 40%. • Set a hard stop-loss: automatically cut one-third of the position if a single asset falls 25%—turning 'I'll cut my losses' from a slogan into an automatic rule. • Do an honest review every three months—does your 'win-rate estimate' still hold?

Are you willing to take this advice? Or did you just come looking for an 'authority' to agree with something you'd already decided?"

Old Wang was stunned again.

That last sentence—"Or did you just come looking for an 'authority' to agree with something you'd already decided?"—hit him with precision.

He had indeed already decided to add to his position. He had come to Claude in order for Claude to help him prove that "this decision was reasonable"—not for Claude to actually help him think.

He shut down the computer.

That day he did not add to his position.

A month later—in April—BTC fell from $88,500 to $60,000—a drop of 32%. Had he increased to 50%, he would have lost 160,000 RMB (about $50,000).

He sent Claude a message:

"Thank you for not letting me add to my position that day."

Claude replied:

"It wasn't that I didn't let you add to your position; it was that you yourself decided not to add. You just needed a 'voice of dissent' to help you see yourself clearly. This is the most important use of the Kelly formula in the age of AI."

13.2 The Paradigm Shift of the Kelly Formula in the Age of AI

This story—though fictional—represents the paradigm shift that the large language models of 2024–2026 have brought to the Kelly formula.

The stories we watched across the first 12 chapters of this book—Thorp, PNP, Bill Gross, Buffett, Bill Benter, Walters, Voulgaris—what they had in common was this: Every one of them had an "internal dissenter"—whether grounded in a mathematical model, in disciplined training, or in the pain of past failures—every one of them could force himself to oppose his own urge to bet.

This is the scarcest ability among professional investors—and most retail investors cannot do it.

And the revolution of the age of AI is this:

A large language model can serve as your "internal dissenter"—it does not have your emotions, does not have your overconfidence, does not have your "this time is different" illusion. Its job is to raise objections to every one of your decisions.

This is not AI making decisions for you—it is AI helping you see yourself clearly.

13.3 The "Workflow" of an AI-Assisted Kelly System

The hands-on quantitative investors of 2024–2026 have already begun integrating AI into the application flow of the Kelly formula. A typical "AI-assisted Kelly system" looks like this:

Step 1: Signal generation

A human (or a quantitative model) proposes an investment idea—"I want to add to my BTC position."

Step 2: Parameter estimation (the three-estimate method)

For each key parameter, the AI has the human give three estimates:

• Optimistic estimate: the most favorable case • Neutral estimate: the most likely case • Pessimistic estimate: the most unfavorable case

For example: the estimate for BTC's μ over the next 12 months—optimistic 30%, neutral 15%, pessimistic 5%.

Step 3: The AI integrates estimates from different models

The AI calls on multiple independent models (such as different LLMs, different quantitative models, different historical windows) to produce independent estimates of the same parameter. Then it integrates them (ensemble)—usually taking the most pessimistic one as the execution parameter.

Step 4: Generating three sets of Kelly fractions

Based on the optimistic/neutral/pessimistic sets of parameters, it computes three Kelly fractions—usually using the pessimistic set.

Step 5: The AI writes a "letter of dissent" (pre-mortem) The AI writes a "letter of dissent"—supposing this trade fails completely, what would happen?

• A list of the possible causes of failure • An estimate of the magnitude of loss upon failure • The trigger conditions for failure • Similar historical cases (how others failed in similar situations before)

This kind of "reverse hypothesis testing" has been shown in cognitive science to significantly reduce decision bias—it forces the human to face his own blind spots.

Step 6: Hard-rule check

No matter how the AI computes, a few hard rules cannot be crossed:

• A single position ≤ 10% of total capital • A single asset class ≤ 30% of total capital • Total leverage ≤ 2x • Distance to liquidation ≥ 5x daily volatility

Any trade that exceeds a hard rule is rejected outright—no matter what the Kelly formula says.

Step 7: Human second confirmation

Any position ≥ 0.3 Kelly compels a human second review—read the letter of dissent, check the hard rules, confirm "I really do want to do this"—and only then can it be executed.

Step 8: Execution

Bet according to the Kelly fraction under the pessimistic parameters.

Step 9: Real-time Bayesian updating

After the trade, use the outcome to update your estimates of the parameters —this is the Kelly formula's "learning loop."

13.4 A Real Comparison

Let us look at a set of real data—a comparison of a semi-professional trader over the 6 months before and after going live with an "AI-assisted Kelly system":

Metric Before After Improvement Monthly trades 42 17 −60% True win rate 51.8% 58.0% +6.2 percentage points Average position 8% 4% −50% Largest single loss −12% −4% −67% Annualized return 8% 14% +6 percentage points Maximum drawdown −22% −8% −64% Sharpe ratio 0.6 1.4 +130%

Key findings:

• The process caused 60% of the trades originally intended to be rejected—most of these, in hindsight, were "losing trades." • The win rate of the remaining 40% of trades improved markedly—because only the highest-quality ideas passed the AI's review. • Total return actually rose—fewer trades, smaller positions, but a higher win rate.

This is precisely what the Kelly formula has been telling us across seventy years—

Real profit comes from "avoiding losses"—not from "seizing every opportunity."

And the greatest value of AI—is that it makes "avoiding losses" executable even for ordinary people.

13.5 The "Five Iron Rules" of the Kelly Formula in the Age of AI

After being tested in practice from 2023 to 2026, the "AI-assisted Kelly system" has formed five new iron rules:

Iron Rule 1: The AI computes Kelly, the human decides whether to execute—do not outsource decision-making authority

What the AI gives is a computational recommendation—but the final "whether or not to press the confirm button" must be done by the human. The human knows his own true risk tolerance, family pressures, psychological state—the AI does not.

Iron Rule 2: Any position ≥ 0.5 Kelly compels a human second review

A bet exceeding half Kelly is always dangerous—at the very least, before execution, force another person to look it over (family, a friend, another investor), and hear one dissenting opinion.

Iron Rule 3: The AI must give an uncertainty interval for "how certain I am"

Have the AI output a "95% confidence interval"—for example, "BTC's expected return over the next 12 months is 15%, 95% confidence interval [−30%, +60%]." This interval is the key to gauging the reliability of the AI's estimate. The wider the interval, the more conservative a Kelly fraction you should use.

Iron Rule 4: The AI's input parameters must be traceable to raw data Do not let the AI give numbers "by feel." Every parameter estimate must have data support—historical samples, market consensus, verifiable facts. Untraceable numbers = untrustworthy decisions.

Iron Rule 5: The AI's letters of dissent must be recorded and archived

Every trade's "letter of dissent"—the AI's writing, the human's response, the final decision—must be archived. This is material for future review—and also a library of material for your long-term learning.

13.6 The Greatest Trap of the Kelly Formula in the Age of AI

But AI has also brought new traps—traps that did not exist in any of the first 12 chapters:

Trap 1: Treating AI as a "predictor"

Many people use AI this way: "Claude, tell me the price of BTC next year."—treating the AI as an oracle.

This is a completely wrong use. The AI's ability to "predict" future prices is the same as a human's—very limited.

The AI's true value is not in prediction—it is in systematic dissent. Having the AI put forward the arguments against a decision you are already inclined toward—this is the AI's core use.

Trap 2: Letting the AI execute directly

Some "AI quantitative trading systems" claim they can be "fully automated"—you don't need to make decisions, the AI automatically judges, automatically places orders, automatically stops losses.

This is dangerous. The AI does not know:

• The true risk tolerance of your account • Your family's financial situation • Your psychological state • Your goals in life

A fully automated AI system may perform very well in a "normal market"—but in moments of crisis it will make catastrophic decisions. The human's final veto power cannot be surrendered.

Trap 3: Misreading the AI's "I don't know" as "no signal"

The AI will often say "I cannot give a reliable estimate of this"—for instance, about certain extreme scenarios, newly emerged assets, or unusual market conditions, it truly does not know. The correct response is: lower the position—high uncertainty = a low fraction recommended by the Kelly formula. The wrong response is: to treat "the AI doesn't know" as "there is no signal" = betting at your usual position—this is overbetting.

The two must not be confused.

Trap 4: Over-reliance on the LLM's "backtests"

An LLM's training data may contain future information—when you ask it to "backtest a certain strategy's performance in 2022," it may unwittingly use

2023 data to 'predict' the 2022 market. This is a look-ahead bias peculiar to the age of LLMs.

The correct approach: run important backtests with a dedicated backtesting engine—do not rely entirely on the LLM.

Trap 5: Ignoring "model correlation"

The outputs of two LLMs (such as Claude and GPT) are highly correlated—they are influenced by the same training data and similar architectures. Having two highly correlated LLMs give "independent estimates"—is not truly independent at all.

A true "model ensemble" must use: LLM + traditional quantitative model + human expert—three genuinely independent sources.

13.7 A Conversation About the Future

Let me end this chapter with a conversation that is fictional but close to reality.

An evening in 2026. A twenty-eight-year-old young investor encounters the Kelly formula for the first time—he has finished reading the first 12 chapters of this book, and has found Claude to chat with.

Him: Claude, I've finished reading this book. I feel the Kelly formula is impressive—but I don't know how to begin.

Claude: First tell me—how much money is in your account? How long do you plan to invest? What is the maximum drawdown you can bear?

Him: 300,000 RMB, 10 years, at most a 30% drawdown.

Claude: Good. Based on this information—

First: your "risk budget" is 30% × 300,000 = 90,000 RMB. Any trade that brings a single position close to this number is overbetting.

Second: your "Kelly fraction" should be quarter Kelly—based on your tolerance for a 30% maximum drawdown, CRRA $\gamma \approx 4$.

Third: you should avoid using any leverage above 2x—crypto perpetual contracts at most 5x. Fourth: your ceilings for a single asset: stocks 25%, crypto 15%, a single bond 5%.

Fifth: your cash reserve should never fall below 30%—these are the bullets kept for "opportunity."

These are the baselines. Within the baselines, make your specific decisions using the "AI-assisted Kelly system" we talked about today.

Him: It sounds so conservative. Can I be a bit more aggressive—say, use half Kelly, single crypto 30%?

Claude: You can—but you shouldn't.

The reason is this—you've only just begun using the Kelly formula. You haven't yet been through any major failure—you don't know your own true risk tolerance.

Buffett, Thorp, Bill Gross, Bill Benter—all of them, when they were young, went through the process of wiping themselves out—and only then did they learn to be conservative.

You first use quarter Kelly—go through a few years of the market's ups and downs—see your own true performance—and then, based on real data, decide whether to adjust.

Remember the deepest lesson of the Kelly formula—

"Admitting what you do not know" is the entry threshold of the Kelly formula.

Right now you do not know yourself—there is no shame in that. But pretending you know yourself—that is where the danger begins.

Him: …Thank you.

13.8 The True Thesis of This Chapter—and of This Book

We have completed the journey of 13 chapters—from a paper at Bell Labs in 1956 whose title was changed, to an AI conversation with a Shanghai retail investor in 2026—

Over seventy years, this formula has traveled to more places than any person:

• It was in a paper at Bell Labs • It was at the gambling tables of Las Vegas • It was in PNP's convertible-bond arbitrage • It was in the two trillion dollars managed by PIMCO • It was in Berkshire's concentrated holdings • It was in Bill Benter's Hong Kong horse-racing algorithm • It was in the discipline of professional sports-betting syndicates • It was in the reflection after every crash in the crypto market • It was in Samuelson's monosyllabic paper • It was in the tragedy of LTCM's collapse, where it was not used • It was in the AI conversations of 2026

The common theme of all these scenes—is a profound thesis of this formula:

The Kelly formula is not a tool of prediction. The Kelly formula is a discipline of allocation.

It does not tell you what will happen in the future. It tells you—given your uncertainty—how you should allocate your capital so as to have the best chance of surviving.

Its most profound insight is not "how to win more"—but rather:

You must honestly know just how much you actually know.

If you are honest—the Kelly formula will give you long-term compounding. If you are not honest—the Kelly formula (or a strategy that does not use Kelly) will take you to zero at some point in time.

Chapter Summary: The Most Profound Insight of the Kelly Formula in the Age of AI

First: AI is not a predictor, AI is a dissenter—its greatest value is helping you oppose your own overconfidence.

Second: AI helps you carry out "reverse hypothesis testing" (pre-mortem)—forcing you to face your own blind spots.

Third: the human retains the final veto power—any fully automated AI Kelly system is dangerous.

Fourth: all the lessons of the Kelly formula still hold in the age of AI—fractional Kelly, pessimistic estimates, single-position ceilings, correlation adjustments, long-term compounding.

The next passage—is the final chapter of this book.

Epilogue · The Direction of Compounding Matters More Than Its Speed

Final.1 A Simple Piece of Arithmetic

Let me begin the ending of this book with the simplest piece of arithmetic.

Suppose there are two investors, each starting with 1 million in capital.

Investor A: earns a steady 1% every day. Investor B: gains or loses 5% every day, at 50/50 odds. After 365 days, how much wealth will each of them have?

Investor A: - 365 days × 1% compounded daily = (1.01)^365 - = 37.78 times over — that is, 1 million turns into 37.78 million.

Investor B: - In expectation, each day 0.5 × (+5%) + 0.5 × (-5%) = 0% - The arithmetic mean tells you "over the long run, no gain and no loss" - But the geometric mean tells you: (1.05 × 0.95)^(365/2) = (0.9975)^182.5 ≈ 0.633 - that is, 1 million turns into 633 thousand — a loss of 37%.

The "daily expectation" of the two men is almost identical — A gains +1% each day, B averages 0% each day.

But after 365 days, A has 60 times what B has.

This is the very same sentence the Kelly formula has been repeating for seventy years —

Compounding is not about how fast you make money. Compounding is about whether or not you can avoid going to zero.

Final.2 What the Kelly Formula Is Really Saying

I have used an entire book to tell 13 stories —

• Kelly wrote down half a line at Bell Labs, died at 41, and never saw the day his formula was put to use • Shannon handed Kelly's paper to Thorp, and himself, in the Kelly spirit, earned 28% annualized over 30 years • Thorp doubled his money in 72 hours in Las Vegas, and then went 19 years on Wall Street without a losing year • Bill Gross went from 200 dollars to 2 trillion in assets under management, carried through 44 years by his "2% rule" • Buffett, with "high conviction + large position," grew Berkshire 50,000-fold over 60 years • Bill Benter, with a model of 120 variables + the Kelly formula, walked away with 1 billion dollars from the Hong Kong racetrack • Walters and Voulgaris survived 30 years in sports betting with a 2% cap per wager • Retail traders in the crypto market, not understanding the Kelly formula, evaporated 5 trillion dollars • Samuelson questioned the Kelly formula with a paper written in single syllables • LTCM, by violating the Kelly formula, lost 90% in 3 months and nearly triggered a global financial collapse • In the age of AI, the Kelly formula became a tool for "arguing against your own impulses"

The shared theme of these stories — there is only one:

The Kelly formula does not give you a precise answer to "how big should I bet." The Kelly formula teaches you a discipline of "being honest with yourself."

All of its mathematical conclusions — full Kelly, half Kelly, quarter Kelly, the CRRA spectrum, fractional Kelly, a stress-tested Σ — are different expressions of this discipline. Its core spirit can be distilled into three sentences that need no mathematics:

First sentence: You can win, because you know what you know.

Second sentence: You can avoid going bankrupt, because you know what you do not know.

Third sentence: You can survive, because you do not confuse these two things.

Final.3 The "Seventy-Year Course" of the Kelly Formula

Let me make one final summary of everything this formula has taught over seventy years:

On winning:

1. Find the real edge — win rate ≠ confidence; it is a genuine advantage based on objective data
2. The bigger the edge, the bigger the bet — this is the core of the Kelly formula
3. When there is no edge, do not bet — this is the often overlooked but equally important side of the Kelly formula

On surviving:

4. Never use full Kelly — use 1/4 to 1/2 Kelly
5. Always control the single position — no more than 2-5% of total capital

6. Always control total leverage — no more than 2-3 times (unless you have an exceptionally strong edge)
7. Always keep cash on hand — over 30%, saving ammunition for "opportunities"
8. Always run the "reverse hypothesis test" — if this one bet is completely wrong, what happens?

On growth:

9. Review honestly — real win rate vs. self-assessed win rate, whether the edge still exists
10. Stop immediately when the edge disappears — better to stop before the edge is completely gone
11. The law of large numbers requires repetition — one large bet is never as good as many small bets

On human nature:

12. Overconfidence is the greatest enemy of the Kelly formula — all "overbetting" springs from it
13. Admitting "I don't know" is the entry threshold of the Kelly formula — those who cannot do it will never use it well
14. Discipline is a skill — once you abandon discipline, no judgment, however clever, can save you

On the age of AI:

15. Let AI be your "opponent," not your predictor
16. Integrate multiple independent sources — one AI's estimate = one estimate; only three independent estimates are robust

17. Humans keep the final veto — never outsource the decision

Final.4 A Final Request to All Readers

Writing to this point — the last few hundred words of this book — I would like to make a request.

Whether you are a retail investor hearing about the Kelly formula for the first time, or a quantitative analyst who has used it for many years —

Please do one simple thing:

Open your account. Look at how large your single biggest position is.

Then ask yourself three questions:

1. For this one bet, can I objectively quantify my win rate? (Not "I feel it will go up" — but "what data do I have to support this estimate")
2. If this one bet is completely wrong, how much will I lose? What impact will my total capital take?
3. If this one bet loses 50%, what will I do? What is my "cut-loss rule"?

If for any one of these three questions you do not have a clear answer — then this one bet of yours has already violated the spirit of the Kelly formula.

No matter what it is called, no matter how bullish you are, no matter what the market says —

Please consider reducing the position.

Not because you are certain to lose — you may well make a great deal. But because winning one bet does not matter — surviving over the long run is what matters.

Final.5 The Direction of Compounding

Having finished this book, I am reminded of a small thing.

Many years ago, in a bookstore, I came upon an old book — its cover tattered, priced at 5 yuan. Its title was Fortune's Formula, by William Poundstone.

That book took me three evenings to finish. After finishing it, I felt differently about everything.

I thought of all the "I-thought-I-was-clever" decisions I had made — which were in fact gambles that rode on luck. I thought of all the trades I believed I had "figured out clearly" — which were in fact built on overconfidence. I thought of the things that had cost me several months' salary in losses — had they happened under the discipline of the Kelly formula, they would at most have cost me a week's salary. The Kelly formula cannot let me earn more — it can let me lose less.

And losing less — after several years of compounding — is far more effective than "earning more."

If you have finished this book, remembering three things is enough —

The first: It was 1956. A 32-year-old physicist wrote down half a line on a notepad at Bell Labs — he would later die of a stroke at 41, having never used the formula he invented.

The second: This formula has been verified countless times, doubted countless times, saved people countless times, and ignored countless times.

Everyone who truly used it survived — Thorp, Bill Gross, Buffett, Bill Benter — and their wealth came not from cleverness, but from discipline.

The third:

The direction of compounding matters more than its speed.

Seventy years ago at Murray Hill, New Jersey, no one would have said this sentence. Today, here, I say it once. Not because I have understood it — but because I have finally stopped pretending to understand it.

May you place your bets honestly. May the direction of your compounding be forever upward. Spring 2026, written in Singapore.

Appendix A · The Kelly Formula Cheat-Sheet

A.1 Basic Formulas

Binary Kelly (basic form)

Win probability p, net odds b, loss probability q = 1 − p. Betting fraction:

b·p − q

f* =

b

Even-odds simplification (b = 1): f* = 2p − 1

Win probability p f* (b=1) half-Kelly 1/4 Kelly

0.51 2% 1% 0.5%

0.55 10% 5% 2.5%

0.60 20% 10% 5%

0.65 30% 15% 7.5%

0.70 40% 20% 10%

A.2 Continuous-Time Kelly (stocks, futures, crypto)

μ − r

$$f^* =$$

σ^2

where: μ = annualized expected return, σ = annualized volatility, r = risk-free rate

Continuous Kelly under typical parameters:

Asset μ σ r π* (full Kelly) 1/4 Kelly S&P 500 10% 16% 3% 273% 68% CSI 300 8% 22% 2% 124% 31% BTC 25% 70% 2% 47% 12%

A.3 CRRA Risk-Aversion Adjustment

μ − r

$$f^* =$$

$\gamma\sigma^2$

γ Name Kelly fraction Who it suits 1 log utility full Kelly theoretical investors 2 moderate aversion half-Kelly most hedge funds 4 high aversion 1/4 Kelly most retail investors, pension funds 10 extreme aversion 1/10 Kelly those near retirement, conservatives

A.4 The "Magic Table" of Fractional Kelly α (Kelly fraction) geometric growth retained volatility (relative to full Kelly) long-run probability of not going broke

1.0 100% 100% ~50%

0.75 93.75% 75% ~80%

0.5 (half-Kelly) 75% 50% ~90%

0.25 (1/4 Kelly) 43.75% 25% >99%

Key observation: half-Kelly sacrifices only 25% of geometric growth, but halves volatility and drops the probability of ruin from 50% to 10% — a very worthwhile trade-off.

Appendix B · A Seventy-Year Timeline

Year Event Protagonist 1948 Shannon publishes "A Mathematical Theory of Communication" — founding information theory Shannon 1953 John L. Kelly Jr. earns his PhD in physics Kelly 1955 Kelly joins Bell Labs Kelly 1956 Kelly publishes "A New Interpretation of Information Rate" Kelly 1959 Latané publishes his geometric-mean-maximization paper Latané

1960.11 Shannon hands Kelly's paper to Thorp Shannon →

Thorp 1961 Breiman proves the long-run dominance of the Kelly strategy Breiman 1961 Thorp + Kimmel play blackjack for real in Reno Thorp 1962 Thorp publishes Beat the Dealer Thorp

1965.3 Kelly dies of a stroke, aged 41 Kelly

1966 Bill Gross plays blackjack in Las Vegas ($200 → $10,000) Gross 1966 Thorp coins the name "Kelly Criterion" in the revised edition of Beat the Dealer Thorp 1967 Thorp publishes Beat the Market, independently guessing an approximation of the Black-Scholes formula Thorp

1969.11 Thorp + Regan found Convertible Hedge Associates Thorp

(later renamed PNP) 1969-79 Samuelson publishes three times against the Kelly formula Samuelson 1971 PIMCO is founded; Bill Gross is on the founding team Gross 1971 Merton derives the continuous-time Kelly formula Merton 1973 The Black-Scholes option-pricing formula is formally published Black, Scholes 1979 Samuelson publishes his monosyllabic anti-Kelly paper Samuelson 1984 Bill Benter + Alan Woods move to Hong Kong Benter 1987 Benter turns his first profit at horse racing, $600,000; the PIMCO Total Return Fund is founded Benter, Gross 1988 Buffett begins buying Coca-Cola in large quantities Buffett

1988.12 PNP is wound down because of an FBI investigation (19 years with no losing quarter) Thorp

1992 MacLean-Thorp-Ziemba publish "Growth versus Security," formalizing fractional Kelly Thorp 1994 LTCM is founded, with 30x leverage Meriwether et al. 1994 Thorp founds Ridgeline Partners Thorp 1997 Scholes and Merton win the Nobel Prize in Economics Scholes, Merton 1998.8- LTCM loses 90% in three months; the Fed organizes a bailout LTCM 9 2001 Bill Benter retires (cumulative profit of $1 billion) Benter

2001.2 Shannon dies Shannon

2002 Thorp voluntarily shuts down Ridgeline Thorp 2007 Pabrai publishes The Dhandho Investor, bringing the Kelly formula into value-investing circles Pabrai 2009 Samuelson dies Samuelson

2014.9 Bill Gross leaves PIMCO Gross

2016 Berkshire begins buying Apple in large quantities Buffett 2018 Sun & Boyd publish their distributionally-robust Kelly paper Sun, Boyd

2020.3 "Black Thursday" — the crypto market sees $3-4 billion of liquidations in a single day —

2021.5 The "5.19" crypto liquidation — $9 billion liquidated in a single day —

2022.5 LUNA collapses — $40 billion of market cap evaporates in 72 hours —

2022.11 FTX fails — $10-15 billion of customer assets frozen SBF

2023-26 AI large language models are integrated into Kelly systems; the "objection letter" process is standardized — 2026 This book is published —

Appendix C · Recommended Reading

Introductory (suitable for all readers):

1. Fortune's Formula by William Poundstone (2005) — the most authoritative popular account of the Kelly formula's 70-year history. Key chapters: the stories of Kelly, Shannon, and Thorp
2. A Man for All Markets by Edward O. Thorp (2017) — Thorp's own autobiography, from blackjack to PNP to Ridgeline
3. Beat the Dealer by Edward O. Thorp (1962/1966) — the classic blackjack textbook

Advanced (suitable for those with a finance background):

4. When Genius Failed by Roger Lowenstein (2000) — the most detailed account of LTCM's 1998 collapse
5. The Mathematics of Poker by Bill Chen & Jerrod Ankenman (2006) — the engineered application of poker and the Kelly formula
6. The Dhandho Investor by Mohnish Pabrai (2007) — the value-investing version of Kelly

Professional (quant researchers):

7. The Kelly Capital Growth Investment Criterion edited by MacLean, Thorp, Ziemba (2011) — the most complete collection of academic papers on the Kelly formula

8. Kelly, J. L., "A New Interpretation of Information Rate" (1956) — the original paper, essential reading
9. Breiman, L., "Optimal Gambling Systems for Favorable Games" (1961) — the rigorous mathematical proof
10. Sun & Boyd, "Distributionally Robust Kelly Gambling" (2018, arXiv:1812.10371) — distributionally-robust Kelly for the AI era

Afterword

Having finished writing this book, I feel a state that is at once weary and relieved.

Weary, because this has been my deepest attempt yet at writing about the Kelly formula — drawing throughout on Kelly's original 1956 paper, Thorp's three books, Poundstone's Fortune's Formula, the investment history of PIMCO, Buffett's letters to shareholders, interviews with Bill Benter, the reports on LTCM's collapse, Samuelson's three dissenting papers, and the latest research on AI-assisted Kelly systems.

Relieved, because I can at last tell the full story of this formula's seventy years — in a way an ordinary reader can actually get through.

I hope this book accomplishes three things:

First: that it makes you remember a few people — Kelly, Shannon, Thorp, Gross, Buffett, Benter, Walters — the real protagonists of this formula's seventy years. Second: that it makes you understand a kind of discipline — the Kelly formula is not a "formula for winning," it is a "formula for not losing." Third: that it makes you change a behavior — the next time you make an investment decision, at least pause for 30 seconds and ask yourself: "Have I really estimated the win probability of this bet objectively?"

If you can manage that third point — then this book will have been worth it.

May the wisdom of the Kelly formula's seventy years accompany you through every trade to come.

May the direction of your compounding always be upward.

The End.

f*

The Wisdom of Betting · Seventy Years of the Kelly Formula

by Robert B.

English edition · F★ Protocol

Not investment advice. Past performance does not predict future results.