

概率：与不确定共处

从三条公理到凯利判据，一门讲给理工读者的概率课

王建硕

费曼式写法 · 由多个 AI 代理撰写与互相审校

导读

从三条公理到凯利判据，一门讲给理工读者的概率课

先别背公式。先回答一个问题：

天气预报说「明天降水概率 70%」。明天只会来一次——这个 70%，到底挂在什么东西上？

如果你卡住了，恭喜，你已经在认真对待概率了。这个问题让数学家吵了两百年：有人说概率是世界的属性（下雨有 70% 的「倾向」），有人说它是频率的影子（相似的天气里七成下了雨），有人说它根本不在世界里，而在你脑中（你愿意用 7:3 的赔率赌它下）。

这本书不打算和稀泥。我们要把概率拆成五块来看：

- **它怎么被定义**——为什么一个词需要三张脸，以及柯尔莫哥洛夫如何用三条公理终结争吵；
- **它数学上是什么**——一个给事件打分的函数，和从三条公理里长出的整套算术；
- **金融界怎么看它**——交易员不问「概率是什么」，他们问「概率藏在哪里」：藏在赔率里、价格里、以及正态分布那张温柔谎言的裂缝里；
- **如何评估它**——概率不是玄学的品位，是可以打分、可以练习的手艺；有人为此办过预测锦标赛，赢家的方法朴素得让人失望；
- **如何靠它做决定**——期望值为什么不够，为什么「全押期望最高的选项」是一条数学上保证通往归零的路。

贯穿全书的问题只有一个：**当未来不确定时，一个理性的人该怎么想、怎么下注、怎么活？**

读法上有一个承诺和一个要求。承诺是：每个名词第一次出现都会当场用人话解释，没有一个公式需要你背。要求是：每章结尾留下的问题，值得你先停一分钟想自己的答案——概率这东西，读是读不会的，要在具体的赌局、体检报告和市场报价里练。

我们从那三张脸开始。

第一章 · 一个词的三张脸

天气预报员到底在说什么

天气预报说：「明天降水概率 70%。」这句话你听过几千遍，但如果有人拦住你问：70% 到底是什么意思？大多数人会卡住。

是说「明天有 70% 的地区会下雨」？不对，那是降水覆盖率。是说「明天有 70% 的时间在下雨」？也不对。气象台的实际意思是：在**和明天各方面条件都相似的日子里**，历史上大约七成下了雨。可问题是——明天只来一次，不会重复一百次让你数。一个只发生一次的事件，它的「概率」究竟挂在什么东西上？

这不是抬杠。这个问题，数学家、统计学家和哲学家吵了两百多年，吵出了三张完全不同的脸。

第一张脸：对称

最古老的定义来自赌场。掷一颗均匀的骰子，六个面长得一模一样，没有任何理由偏爱哪一个，于是每个面出现的概率是 $1/6$ 。这叫古典概率——大白话就是：**数出所有「看起来一样」的可能结果，favorable 的占几份，概率就是几分之几。**

拉普拉斯在 1814 年把这个想法写成了经典定义。它漂亮、干脆，做排列组合题所向披靡。但它的地基藏着一个循环：「六个面一样可能」——「一样可能」是什么意思？就是「概率相等」。用概率定义概率，绕了个圈。

更深的裂缝是：一旦结果不对称，这张脸就不认识了。一枚被做了手脚的骰子、一场球赛的胜负、明天是否下雨——没有一个天然的「等可能格子」让你数。

第二张脸：频率

那就别空想，做实验。掷一万次骰子，数出「6」出现了多少次，次数除以一万，就是概率。掷的次数越多，这个比例越稳定——这叫频率概率，大白话：**概率就是一件事在无穷多次重复中的出现频率。**

这张脸踏实、可测量，是现代统计学（假设检验、置信区间那一整套）的地基。保险公司用它算保费：不用懂你为什么出事，只要知道去年十万司机里出了多少起事故。

但它有一个致命的尴尬：「无穷多次重复」在现实中不存在。你永远只能掷有限次，得到的只是概率的近似。更麻烦的是那些只发生一次的事——2026 年这场比赛、这次手术、这笔投资——没有「重复一万次」可言。频率学派面对单次事件，只能含糊地说「概率不适用于单个事件」。可我们偏偏最想给单次事件一个数。

第三张脸：信念

第三条路最激进：概率根本不是世界的属性，而是**你脑中的东西**。说「明天下雨概率 70%」，不是说世界有 70% 在下雨，而是说：以你掌握的信息，你愿意用 7:3 的赔率赌它下雨。这叫主观概率（也常叫贝叶斯概率），大白话：**概率是量化后的信念强度，随信息更新**。

意大利数学家德·菲内蒂给了一个漂亮的操作化：别空谈信念，报价。你声称「明天下雨概率 70%」，那我设计一个赌局——你 70 块钱买一张「下雨就兑 100」的票，你觉得公平吗？如果公平，你的概率就是 0.7，否则你在撒谎。信念被迫用钱显形。

大白话

大白话：古典派数格子，频率派数次数，主观派数你愿下多大的注。前两张脸把概率当成**世界的属性**，第三张脸把它当成**人的属性**。

三张脸谁对？

都对，各自有地盘。扑克牌和骰子是古典派的天下；工厂质检、药物试验是频率派的天下；而「这次创业能不能成」「这支股票会不会涨」——没有重复、没有对称——只有主观派接得住。

更有意思的是，三张脸共享同一套算术：都在 0 到 1 之间，互斥事件可以相加，该乘的时候乘。1933 年，柯尔莫哥洛夫看穿了这一点：别再争概率「是」什么了，只管它**怎么算**。他用三条公理把三种解释统一在同一个数学框架下——那是下一章的故事。

先记住这本书的立场：我们把概率当作工具而不是本体。它能不能帮你在信息不全时下一个判断、做好一个决定，才是我们关心的事。至于它到底是世界的皱纹还是你脑中的刻度——你会看到，真正有趣的是，这个区分在很多场合根本不影响你怎么行动。

第二章 · 三条公理撑起一切

不争「是什么」，只管「怎么算」

上一章的三张脸吵了两百年。1933 年，31 岁的苏联数学家柯尔莫哥洛夫（Andrey Kolmogorov）用一本不到一百页的小册子终结了争吵。他的招数是数学家的经典太极拳：**拒绝回答概率是什么，只规定概率满足什么。**

就像几何学不定义「点是什么」，只说点、线、面满足哪些关系，剩下的全靠推导。公理化的妙处在于：古典派、频率派、主观派——只要你的用法满足这几条规则，后面的整套数学你就随使用，管你心里怎么解释。

先把舞台搭好

公理之前需要两个道具。样本空间：所有可能结果的清单，大白话就是「事情能怎么发展，先全部列出来」。掷骰子的样本空间是 $\{1,2,3,4,5,6\}$ 。事件：样本空间的一些子集，即「你关心的那几种发展」。「掷出偶数」就是事件 $\{2,4,6\}$ 。

概率，就是一个给每个事件打分的函数 P ：输入一个事件，输出一个数。柯尔莫哥洛夫说，这个函数只要满足三条：

1. **非负**：任何事件的概率 ≥ 0 。没有负的可能性。
2. **归一**：整个样本空间的概率 $= 1$ 。「总会发生点什么」，是确定无疑的。
3. **可列可加**：如果一堆事件两两互斥（不能同时发生），那么「其中任意一个发生」的概率，等于各自概率相加。掷出 2 或 4 的概率 $= P(2) + P(4)$ 。

就这三条。中学课本里所有概率公式——加法公式、乘法公式、全概率、贝叶斯——全是这三条的推论。

三条公理偷偷放进来了什么

看几条立刻能推出来的东西：

- **补集**： $P(\text{不发生}) = 1 - P(\text{发生})$ 。因为「发生」和「不发生」互斥且合起来是全集，用公理 2 和 3 一步推出。
- **空事件**： $P(\text{不可能}) = 0$ 。「什么都不发生」是「发生点什么」的补集。
- **单调**：如果 A 发生时 B 一定发生（A 是 B 的子集），那么 $P(A) \leq P(B)$ 。更具体的要求不会更可能。

注意最后这条的分量。著名的「琳达问题」里，人们认为「琳达是银行柜员且热衷女权运动」比「琳达是银行柜员」更可能——可前者明明是后者的子集，公理说它**绝不可能更大**。三条干巴巴的公理，直接宣判了一种常见直觉为错误。

「可列」两个字是留给无限的后门

第三条公理里有个容易滑过去的词：可列可加。不是「有限个互斥事件可加」——那太容易了——而是**无穷多个（但能用自然数编号）的互斥事件也能加**。

为什么需要这个？考虑「连续掷硬币直到第一次出正面」这个游戏。可能第 1 次就出，可能第 2 次，可能第 100 万次——这是一个天然的无穷样本空间。「游戏会结束」的概率是 $1/2 + 1/4 + 1/8 + \dots$ ，必须允许无穷项相加，才能算出它等于 1。没有「可列」两个字，概率论连这个小学生游戏都处理不了。

大白话

大白话：公理 3 说「可能性可以加起来」，「可列」是补丁：哪怕可能性分成无穷多份，只要数得清（1、2、3……编号编得完），就还能加。正是这块补丁，让微积分级别的数学得以进入概率论。

代价：有些事件不许问概率

公理化不是白捡的便宜，它有代价。当样本空间是连续的时候——比如「在 $[0,1]$ 区间随机取一个点」——数学家证明了：**不可能给所有子集都分配概率**而不自相矛盾（这牵涉选择公理，维塔利集合是经典反例）。

柯尔莫哥洛夫的解决方案是认怂：不给所有子集打分，只对一篮子「性质良好」的子集（行话叫 σ -代数）打分。这就像测度论里承认「有些集合没有面积」——听着惊悚，实际上你这

辈子关心的所有事件都性质良好。这个限制唯一的实战意义是提醒你：概率论有边界，边界外的问题（比如「从所有整数中等概率随机取一个」——做不到，加起来要么是 0 要么是无穷）问法本身就是错的。

三条公理之后

回头看第一章的争吵：古典概率是公理的特例（每个格子给 $1/n$ ）；频率概率满足公理（频率本来就非负、归一、互斥可加）；主观概率也满足——德·菲内蒂甚至反证过：**如果你的报价违反这三条公理，我就能设计一组赌局让你稳输不赔**（荷兰赌论证）。三条公理是三张脸的最大公约数。

但公理只说了概率的静态规则，还没碰最有戏的问题：当**新信息进来**，概率该怎么变？那是「条件」的世界，下一章见。

第三章 · 信息改变世界

一次体检的惊吓

假设有一种病，在人群中的患病率是 0.1%——一千个人里一个。现在有一种检测，质量很好：有病的人 99% 能测出来（灵敏度 99%），没病的人 99% 能排除（特异度 99%，也就是假阳性率 1%）。

你去体检，结果是**阳性**。你得病的概率是多少？

绝大多数医生面对这个问题会答「99%」或「八成左右」。正确答案大约是 **9%**。这不是脑筋急转弯，是算术。我们把这个算术做一遍，因为这一遍值得记一辈子。

别用百分比，数人头

想象一万人来体检。

- 真有病的人：一万 $\times$ 0.1% = **10 人**。检测灵敏，其中约 9.9 人测出阳性。
- 没病的人：**9990 人**。检测有 1% 假阳性，其中约 99.9 人被误报阳性。

现在看所有拿到阳性报告的人：9.9 + 99.9 $\approx$ 110 人。其中真有病的只有约 10 人。所以你阳性且真有病的概率 $\approx$ 10/110 $\approx$ **9%**。

直觉错在哪？它只盯着「检测有 99% 的准头」，却忘了这场考试里**没病的考生是有病考生的一千倍**——哪怕每个没病的人只有 1% 的概率被冤枉，冤枉者的绝对数量也把真病人淹没了。被忽略的那个 0.1%，叫基准率：一件事在没有任何检测信息之前的「出厂概率」，大白话就是**这事本来有多常见**。

条件概率：把世界缩小

上面的算术背后是一个定义。条件概率 $P(A|B)$ ：在 B 已经发生的条件下，A 发生的概率。大白话：**知道 B 之后，世界缩小了，只在 B 里面重新数 A 占多大比例**。

公式是 $P(A|B) = P(A \text{ 且 } B) / P(B)$ 。别背它，理解它：分子是「既是阳性又有病」那小撮人，分母是「所有阳性」这圈人——样本空间从「所有人」换成了「阳性的人」，重新算比例，仅此而已。条件概率没有引入任何新数学，它只是**换了一个更小的舞台**。

贝叶斯定理：换舞台的公式版

把刚才数人头的过程写成公式，就是贝叶斯定理：

$$P(\text{病}|\text{阳性}) = P(\text{阳性}|\text{病}) \times P(\text{病}) / P(\text{阳性})$$

读法：**看到证据后的信念 = 证据在假设下的可能性 × 原先的基准率 ÷ 证据的总常见度**。

注意右边三样东西的角色：

- $P(\text{病}) = 0.1\%$ ：先验，看证据之前你信多少；
- $P(\text{阳性}|\text{病}) = 99\%$ ：似然，如果假设为真，证据有多可能出现；
- $P(\text{病}|\text{阳性}) \approx 9\%$ ：后验，看完证据你该信多少。

直觉的错误在于把 $P(\text{阳性}|\text{病}) = 99\%$ 直接当成了 $P(\text{病}|\text{阳性})$ 。这两个数差了一个数量级，而它们之间的换算器就是基准率。**贝叶斯定理不是新逻辑，是防止你把这两个条件概率搞混的记账系统。**

这台记账系统到处在跑

- **垃圾邮件过滤**：「免费」一词出现在垃圾邮件里很常见（似然高），但你的老板也可能写「这次团建免费」（基准率捣乱）。贝叶斯过滤器干的正是把两者都记上账。
- **癌症筛查的争议**：对低危人群做筛查，假阳性淹没真病人的算术一模一样——这正是各国对「要不要全民筛查 PSA/钼靶」争论不休的数学内核。
- **法庭上的 DNA**：「匹配概率百万分之一」是 $P(\text{匹配}|\text{无辜})$ ，陪审团需要的却是 $P(\text{无辜}|\text{匹配})$ ——城市里几百万人，百万分之一的匹配会冤枉好几个人。这个错误真实地把人送进过监狱。

大白话：别问「这个检测有多准」，要问「在拿到阳性的人里，真病人占几成」。前一个数是检测的属性，后一个数才是你的处境——而它们之间隔着基准率这条大河。

信息改变了世界吗？

回到本章标题。检测结果没有改变你的**身体**，却把你的概率从 0.1% 推到了 9%——变了九十倍。这泄露了条件概率的本性：概率（至少在本书的用法里）不是世界的状态，而是**你与世界之间信息关系的状态**。同一个你，同一份体检报告，知道基准率的医生和不知道的病人，活在两个概率不同的世界里——而只有一个世界是对的：把账记全的那个。

下一章我们要问一个看似相反的问题：如果信息这么重要，那为什么「已经连续出了五次正面」这条信息，对下一次掷硬币却**一文不值**？

第四章 · 硬币没有记忆，市场有

蒙特卡洛的那个晚上

1913 年 8 月 18 日，蒙特卡洛赌场的一张轮盘桌上，小球连续落在了黑色上。五次、十次、十五次……赌客们疯了一样押红色——「该轮到红了！」——赌注越堆越高。黑色最后连续出现了 **26 次**，赌场当天进账数百万法郎。

赌客们的推理听起来无法反驳：连续出黑的概率是 $(18/37)^{26}$ ，一个天文数字那么小的事件，怎么可能继续？这个推理里藏着一个致命的偷换：从赌局开头看，「连出 26 次黑」确实几乎不可能；但今晚的赌客不是站在开头，他们站在第 15 次黑的现场。**已经发生的 15 次黑是既成事实，概率是 1**，不是天文数字分之一。对下一次旋转而言，条件概率就是平平无奇的 18/37。

独立：信息到此为止

这里登场的概念是独立性：如果 $P(A|B) = P(A)$ ——知道 B 之后，A 的概率纹丝不动——就说 A 和 B 独立。大白话：**B 这条信息对判断 A 一文不值**。用上一章的语言说：世界没有因为新信息而缩小。

硬币和轮盘是独立的教科书标本。物理上讲，小球没有记忆，它不「欠」红色任何东西，也没有「势头」要延续黑色。上一章说信息改变世界，这一章要补上硬币的另一面（双关是有意的）：**有些信息根本不携带信息**。识别「哪些信息是噪声」，和「如何从信息中更新」，是同一种手艺的两半。

这个错误有个名字：赌徒谬误——以为独立序列会「自我纠偏」，短期偏离必须马上被拉回。它错在把大数定律（第六章详谈）错装到了下一次投掷上：大数定律说的是**长期频率的稀释**，不是**下一次的补偿**。硬币还清「欠账」的方式不是多出红，而是让那 26 次黑在上万次旋转中被稀释到无关紧要。

可是——等等，市场真的独立吗？

现在把场景从赌场换到交易所。「股价已经连跌五天了，该反弹了吧？」这句话的结构和蒙特卡洛赌客一模一样。但这一次，它**未必是错的**。

区别在于：轮盘的每次旋转由物理定律保证互不相干，而市场里的每一次「投掷」背后是人——人有记忆、会恐慌、会止损、会追加保证金。连跌五天可能触发的强平链条、恐慌传染、机构风控被动减仓，都会让第六天的分布和第一天**不一样**。反过来，均值回归的力量（估值锚、逢低买盘）又确实存在。市场不是硬币：相关性——两次「投掷」结果之间的统计牵连——在金融数据里真实存在，比如波动率聚类：大涨大跌的日子倾向于扎堆出现，哪怕方向不可预测。

大白话

大白话：「连跌五天该反弹」在轮盘上是谬误，在市场上是一个**可检验的假设**。区别不在句子的语气，在背后的机器：硬币每次重置，市场带着昨天的伤疤和仓位上场。

真正难的问题：这台机器独立吗？

于是概率素养里最难的一步浮出水面：**在使用任何历史数据之前，先判断你面对的过程是不是独立的**。这个判断不能靠数据本身看出来——任何有限序列都能被讲出「规律」。它靠对生成机制的理解：

- 骰子、轮盘、洗牌：物理隔离，独立性可以直接**假设**（除非有人出千）。
- 天气：强相关，明天下雨和今天下雨明显不独立。
- 市场：方向近似独立（所以择时极难），波动强相关（所以风险会抱团）。
- 体育、选秀、面试：人高估连号的「异常」。连续三次硬币正面就让人起疑，其实在随机序列里，连号出现的频率远超直觉——经典课堂实验里，老师让一半学生真掷一百次硬币、另一半凭空编造序列，教过概率论的人一眼就能分开：编出来的那叠，最长的连号明显太短。

一个统一的眼光

把第三、四章合起来，你会得到一个简洁的判别式：拿到一条新信息 B，先问 **$P(A|B)$ 和 $P(A)$ 差多少**。差得多，B 是真信息，做贝叶斯更新；差是零，B 是噪声，扔掉——哪怕它

穿着「连出五次黑」这样戏剧性的外衣。

赌徒谬误的本质，是在独立的世界里强行做更新；而市场上许多人的破产，是在相关的世界里假装独立。两个错误互为镜像，犯错的开关是同一个：没问这台机器有没有记忆。

到此为止我们一直在谈「可能性有多大」。但做决定还需要另一半信息：每种结果值多少。把「多大可能」和「值多少」乘起来，就是下一章的主角——期望值。

第五章 · 把未来折成一个数字

赌局值多少钱？

前两章我们只会回答「可能性多大」。但做决定时你还关心另半边：每种结果带来多少。把两边合起来的装置，是概率论里第一个真正「面向行动」的概念——期望值：每种结果的取值乘以它的概率，再全部相加。大白话：**把未来的所有分支按可能性加权，折成一个平均数**。

一个赌局：掷硬币，正面赢 100 元，反面输 50 元。期望值 $= 0.5 \times 100 + 0.5 \times (-50) = 25$ 元。它的含义是：如果你把这个赌局玩一万次，平均每局净赚 25 元。于是有了决策的第一条黄金律：**正期望值的赌局值得玩，负期望值的不值得**。赌场对你设下的每个游戏期望值都是负的——这就是「久赌必输」的全部数学。

圣彼得堡：期望值撞上现实的裂缝

1738 年，丹尼尔·伯努利在《圣彼得堡科学院评论》上发表了一个让期望值难堪的思想实验（问题由他的堂兄尼古拉·伯努利提出）。赌局规则：连续掷硬币直到第一次出正面，如果第 n 次才出，你赢 2^n 元。第 1 次出正面赢 2 元，第 2 次赢 4 元，第 10 次赢 1024 元……

这个赌局的期望值是多少？

$$E = \frac{1}{2} \times 2 + \frac{1}{4} \times 4 + \frac{1}{8} \times 8 + \dots = 1 + 1 + 1 + \dots = \text{无穷大}$$

每一项都是 1，有无穷多项。按黄金律，你应该愿意付任何价格买这个赌局——一百万也买。但真问你愿意付多少，大多数人的答案是十几二十块。这就是圣彼得堡悖论：期望值说「值无穷」，直觉说「值二十」，两者差了无穷倍。总有一个错了。

伯努利的拆招：钱不是钱

伯努利的解答是经济学史上最富成果的想法之一：错的是「用金额折未来」。**同样一块钱，对不同财富的人价值不同**——对乞丐是半条命，对亿万富翁是找零。他提出真正该被平均的不是钱，而是钱带来的效用：钱对你的主观价值，大白话就是「这笔钱实际改善你生活的程度」。他猜测效用大约是财富的对数：财富翻一倍，效用只加固定的一格。

用对数效用重算圣彼得堡： 2^n 元带来的效用不再随 n 爆炸，期望效用收敛到一个有限的小数——约等于几十块钱。和直觉对上了。悖论消失。

这个修补的影响远超一个赌局。它解释了你为什么买保险（保额期望值为负，但破产的负效用太大）、为什么分散投资（两个不相关的 50/50 项目比全押一个好，钱的期望相同，效用的期望更高）。两百年后冯·诺依曼和摩根斯坦把它公理化：只要你的偏好满足几条自治规则，就存在一个效用函数，让你的选择等于「最大化期望效用」。期望值活下来了，但折的东西从「钱」换成了「效用」。

期望值的盲区：它抹平了分布的形状

就算接受效用修补，期望值作为决策工具仍有一个结构性盲区：**它把所有分支揉成一个数，丢掉了分布的形状**。两个期望值相同的赌局——一个「稳赢 50」，一个「1% 概率赢 5000，99% 归零」——在期望值的眼里是同一个东西，在你的生活里完全是两种人生。

丢掉的信息至少有三块：

- **散布有多大**（方差/波动）：平均赚 25，是波澜不惊还是过山车？
- **最坏会怎样**（左尾）：有没有一个分支是「归零出局」？期望值对「小概率大灾难」的压缩是最危险的——这正是第十二章「毁灭」的主题。
- **不对称**（偏度）：彩票是负期望加正偏度（大概率小亏，极小概率暴富），卖保险是正期望加负偏度（大概率小赚，小概率大亏）。人为什么同时买彩票和保险？期望效用理论答得很吃力，第十一章前景理论会回来收这笔账。

大白话

大白话：期望值像只报平均分不报分布的老师：两个班平均分都是 75，一个班人人 70-80，另一个班一半满分一半不及格。平均分没撒谎，但你要问的下一个问题它没答。

还有一个前提：你能玩很多次

最微妙的一条藏在定义里：「玩一万次，平均每局……」期望值的全部合法性建立在重复上。可人生里最重要的决定——选职业、结婚、创业、押上全部积蓄——都是单次事件，没有一万次平均可言。

单次决策里，期望值还剩多少效力？这个问题会在第六章（大数定律到底保证了什么）和第十二章（为什么全押正期望值的选项也可能通向归零）两次回来咬我们。先把它放在桌上：**期望值是把未来折叠成一个数字的绝佳工具，但折叠必有损失——它默认你能承受「平均」的过程。**

第六章 · 长期会回归，但你未必活得到长期

统计学里最被滥用的「安慰剂」

「长期来看会平均的。」这句话你一定听过：连败之后说「该转运了」、亏损之后说「长期会回来」。它的数学原型是大数定律：同一个实验独立重复很多次，样本平均值会收敛到期望值。大白话：**次数够多，运气会被摊平，平均值露出真身。**

这是统计学全部合法性的来源：为什么抽样一千个选民能推断全国？因为大数定律承诺，样本均值离真值的偏差随样本量缩小。但这条定律的实际内容比它的民间版本窄得多，也怪得多。值得把它拆开看三层。

第一层：它说了什么

数学上有两个版本。**弱大数定律**：样本量 n 足够大时，样本均值偏离期望超过任何小量的概率趋于零。**强大数定律**：不仅「大概率不偏离」，而是偏离这件事几乎必然只发生有限次——均值**几乎必然收敛**。差别微妙（概率收敛 vs 几乎必然收敛），实战中你只需要记住共同的核心：**保证的是「平均值的极限」，不是「下一次的结果」。**

这就是第四章赌徒谬误的判决书：大数定律从不说「连续出了五次黑，下一次更可能红」。它实现收敛的方式是**稀释**——前 26 次黑在十万次旋转里占比趋于零——而不是**补偿**。硬币不欠账。

第二层：「长期」有多长？

大数定律是个极限定理：它承诺 n 趋于无穷时收敛，但对「一万次够不够」保持沉默。收敛速度取决于分布的胖瘦。瘦分布（硬币正反、身高）收敛快，几百次就稳；**胖尾分布收敛慢得残忍**——第八章会看到，金融收益这种分布，「长期」可能长达几百年，远超任何一个人的投资生涯。

凯恩斯那句话放在这里正合适：*长期来看我们都死了*。大数定律说「长期来看平均值是对的」，它没撒谎；只是「长期」这个词在数学里指无穷，在你的生命里指三四十年。两者之

间，隔着所有真实的破产。

第三层：它没保证的那个东西——遍历性

这里藏着全书最重要的一个区分。问：一个正期望值的赌局，值不值得押上全部身家反复玩？看一个具体例子：每局掷硬币，赢了资产 $\times 1.5$ ，输了 $\times 0.6$ 。单局期望 $= 0.5 \times 1.5 + 0.5 \times 0.6 = 1.05$ ，期望收益 $+5\%$ ，正期望。押吗？

算一下连玩十局、运气平平（五赢五输）的结果： $1.5^5 \times 0.6^5 \approx 7.59 \times 0.078 \approx \mathbf{0.59}$ 。赢了期望，输光了钱。再来二十局（十赢十输）： $0.59^2 \approx 0.35$ 。玩得越久，越趋近于零。
期望值是正的，而几乎每个人的结局是归零。

哪里出了鬼？期望值的定义是「所有可能结果的加权平均」——行话叫系综平均：想象一万个平行宇宙里的你各玩一局，取平均。而你实际经历的是时间平均：一个你，一局接一局乘下去。对加法型游戏（每局独立加减固定金额），两者一致；对乘法型游戏（每局按比例缩放你的资产，投资正是如此），两者可以天差地别。当「许多平行宇宙的平均」等于「一个你在时间中的平均」时，这个系统叫遍历的；不相等时——比如上面的乘法赌局——用期望值指导单人决策就是系统性自杀。

大白话

大白话：系综平均问「一万个赌徒平均赚多少」；时间平均问「一个赌徒玩一万局最后剩多少」。乘法世界里这两个答案可以反号：一万人的平均在赚钱，而几乎每一个人都在变穷。平均值对个体撒了谎。

为什么乘法世界这么凶

乘法过程的毒药在于：亏损和盈利不对称。资产跌 40%，需要涨 67% 才能回来；跌 90%，需要涨 900%。一次深亏会把复利曲线的基部削掉，之后所有的「正期望」都作用在一个变小的本金上。时间平均惩罚的不是期望值低，而是波动中的深坑——这正是第十二章凯利判据要优化的东西：不是最大化期望收益，而是最大化时间平均的增长率。

把六章连成一句：**概率论的前半本教你算「平行宇宙的平均」，后半本要教你问「我一个人沿着时间走下去会怎样」**。这两个问题在硬币上重合，在财富上分裂。带着这个分裂，我们

去看一个毕生研究「平行宇宙平均」的行业——金融。

第七章 · 概率藏在价格里

交易员不问哲学

前三章吵的「概率是什么」，在金融从业者那里有个干脆的回答：**不管它是什么，它得定价**。赛马场的赔率、期权的报价、信用违约掉期的点差——金融的第一课是：每一个价格背后都藏着一个概率，会读价格的人能把概率抠出来。

赔率就是穿了马甲的概率

赛马场给一匹马开出 3 赔 1 的赔率：押 1 元，赢了拿回 4 元。这个报价隐含了庄家对这匹马获胜概率的判断——如果真实概率是 p ，押注的期望值是 $4p - 1$ ，赌局「公平」（期望为零）时 $p = 1/4$ 。**赔率的倒数就是隐含概率**。

真实的赌盘会留下水分：把一场比赛所有马的隐含概率加起来，结果超过 100%（比如 115%），超出的部分叫 *overround*，就是庄家的抽头。所以读盘口的第一课：先把所有报价归一化除以 115%，剥掉抽头，剩下的才是市场真正的「集体信念」。

这套读法不限于赌场。预测市场（PredictIt、Polymarket 这类「用真金白银对事件下注」的交易所）里，「某候选人当选」合约卖 0.63 美元，就是市场对当选概率的直接报价：63%。研究显示，这类市场在多数时候比民调更准——原因不是参与者更聪明，而是**说错话要亏钱**，这比任何民调纪律都能过滤胡说。

期权：价格里藏着整条分布

更猛的在期权市场。一份「行权价 100 元的看涨期权」的价格，取决于「到期时股价超过 100 的概率」和「超多少」。妙处在于：不同行权价的期权有一整排报价，理论上可以反推出市场对股价**整条概率分布**的定价——不只是「涨不涨」，还有「涨到每个价位的概率各是多少」。1978 年 Breeden 和 Litzenberger 证明了这个逆向工程在数学上可行。

于是交易员有个独有的武器：拿「价格隐含的概率分布」和「自己估计的真实分布」对比，两边不一致的地方，就是赌注。**金融交易的本质，是对两种概率之间的价差下注**。

但是：价格里的概率是「加过税」的

这里有一个外行几乎必掉进去的坑。期权定价用的那套概率——著名的 Black-Scholes 模型里藏在 $N(d_2)$ 里的概率——**不是真实世界的概率**。它叫风险中性概率：一个假想平行宇宙里的概率，那个宇宙里所有人对风险无所谓，所有资产的期望收益率都等于无风险利率。

为什么要发明这个假宇宙？因为期权的「公平价格」不取决于「涨的概率多大」，而取决于「复制它的成本多高」——用股票和借款动态对冲，可以把期权的未来现金流精确复制出来，复制成本就是唯一不送钱的价格（否则有套利）。这套复制论证里压根没出现真实概率，出现的只是风险中性概率。

大白话

大白话：真实概率回答「这事多可能发生」，风险中性概率回答「市场愿意为这个可能性收多少钱」。两者差一层「风险的价格」——就像一件东西的**使用价值**和**售价**之差。深度虚值看跌期权（灾变保险）的风险中性概率几乎总是高于真实概率，因为持有者愿意为「怕」付溢价——这就是卖保险长期赚钱的来源。

价格投票机的局限

「市场用价格给概率投票」很迷人，但记住两个限度。其一，价格是**边际成交者**的信念，不是全知的神谕——2007 年 CDS（债券违约保险）的价格说次贷债券很安全，它报价的不是概率，是当时愿意卖保险的人的乐观。其二，价格里混着风险溢价、流动性、情绪，把价格当概率读，永远要问一句：我抠出来的这个数，几成是信念，几成是恐惧的税？

下一章看这个税最重的地方：当真实的分布长了一条胖尾巴，而定价模型假装它是正态的时候。

第八章 · 正态分布的温柔谎言

一条优雅曲线的统治

如果概率论有一件镇馆之宝，那一定是正态分布：那条对称的钟形曲线，中间高、两头低，由均值（中心在哪）和标准差（摊得多开）两个参数完全决定。大白话：**大多数取值挤在平均附近，偏离越远越罕见，罕见程度按一个精确的衰减公式暴跌。**

它的统治地位来自一条深刻的定理——中心极限定理：大量**独立的、微小的**随机因素叠加起来，不管每个因素各自是什么分布，总和都趋向正态。大白话：如果一件事是「许多小误差的平均」，那它几乎必然长成钟形。身高是正态的（基因、营养等许多小因素相加），测量误差是正态的。高斯用它处理天文观测误差，它因此也叫高斯分布。

正态分布的「衰减公式」有多狠？偏离均值 6 个标准差（「六西格玛事件」）的概率约为五亿分之一。如果日收益是正态的，1987 年 10 月 19 日那种 -20% 级别的暴跌，按标准差换算需要等宇宙年龄的无数倍才该出现一次。它在一个普通的星期一出现了。

曼德博的棉花

第一个系统性地对这条曲线开炮的是伯努瓦·曼德博（Benoît Mandelbrot）。1963 年他研究棉花价格的百年数据，发现两件事：价格变化的分布比正态「头更肥、尾更粗」；而且大变化和小变化混杂出现的方式，在不同时间尺度上惊人地相似（这个观察后来长成了分形几何）。他主张用肥尾分布（如稳定帕累托分布）替代正态：尾巴衰减慢得多，极端事件不是十亿分之一，而是千分之一、百分之一量级。

塔勒布后来把这个思想推向大众，起了一个传播学上天才的名字：黑天鹅——在正态模型里「不该存在」、在现实中定期造访、造访后还被解释成「早有预兆」的极端事件。但请注意本章标题的立场：**黑天鹅在数学上并不黑**。它只在「用错了分布」这个意义上是黑天鹅——换一个尊重肥尾的模型，它就只是灰天鹅、甚至白天鹅。真正的黑，是模型和现实之间的那道缝。

为什么会肥尾：世界不是独立小因素的叠加

中心极限定理有两个前提：因素要**独立**，单个因素的体量要**小**。金融市场两个都违反：

- **不独立**：第四章说过，市场有记忆。恐慌会传染，止损会连环触发，强平会互相踩踏——一个人的卖出是另一个人卖出的原因，因素之间正相关，叠加定理失效。
- **个体不小**：当一笔主权基金或一个爆仓的对冲基金就能移动市场时，「许多小因素的平均」这个图景本身就垮了。

于是收益分布长出**肥尾**（极端值远比正态预测的多）、**偏态**（跌的时候比涨的时候更陡——崩盘是跳下悬崖，牛市是爬楼梯）、**超额峰度**（同时更爱出小波动和巨波动，「温和的中等波动」反而少）。标准差这把尺子在这种分布上还在，但它量不出真正的风险——就像用平均气温描述一个旱涝交替的地方。

用错模型的代价：LTCM

1998 年的长期资本管理公司（LTCM）是这门课的学费缴纳者。两位诺贝尔经济学奖得主（默顿、斯科尔斯——就是 Black-Scholes 的 S）坐镇，模型精密，前四年年化 40%。他们的策略本质是大量收敛交易：赌价差回归，杠杆加到极高。模型的风险估计基于「收益率近似正态、各国市场大体独立」。1998 年俄罗斯违约，全球市场突然一起动——独立假设崩了，尾部比模型想的粗几个数量级——数周内亏掉 46 亿美元，美联储不得不组织华尔街集体救助，以免连环倒塌。

大白话

大白话：LTCM 不是死于运气差，是死于把「需要宇宙年龄才该发生一次」的事件当成不可能，然后加了 25 倍杠杆去赌它不发生。肥尾世界里，杠杆就是把模型误差放大成死亡的那种东西。

对普通人的三个推论

- **「百年一遇」是话术**。它通常意味着「在正态模型里百年一遇」——在肥尾的现实里可能是十年一遇。听到这个词，先问用的是什麼分布。
- **波动率低 ≠ 安全**。正态框架里风险=波动率；肥尾框架里，最危险的时刻恰恰是长期平静之后（平静把杠杆喂大，把警惕喂小）。塔勒布称之为「火鸡问题」：火鸡的一千天数据都说

主人是朋友。

- **期权价格会出卖肥尾。**第七章说过期权价格隐含分布——真实市场隐含分布的尾巴比 Black-Scholes 的正态尾巴厚（行话叫「波动率微笑」），1987 年之后尤其如此。市场用自己疼过的方式，给正态谎言标了价。

至此我们有了看世界的两只眼睛：概率怎么定义、怎么算（前六章），金融界怎么用价格给概率投票、又怎么在尾巴上摔跤（这两章）。接下来换个主语——不是「世界有多随机」，而是「你估的概率准不准」。

第九章 · 给不确定性打分的人

你说「八成能成」，是哪来的八成？

前面八章把概率当客观对象来研究。现在转个身：当你自己说「这个项目八成能成」「这事我 90% 有把握」时，这些数字配被信任吗？

好消息是，这不是品味问题，有客观的判卷方法。坏消息是，未经训练的人类在这张考卷上普遍不及格——好消息的另一半是，**它可以练，而且练法简单得让人失望。**

校准：说 70% 的事，就该七成发生

评判概率的第一把尺子叫校准：把你所有「70% 把握」的预测收集起来，其中应该约 70% 真的发生；「90% 把握」的应约 90% 发生。大白话：**你的信心刻度得和现实的命中率对得上。**

经典的失败样本是普通人做常识填空题：「我有 90% 把握尼罗河比长江长」——让受试者给几百道题标信心，结果「90% 有把握」的答案只对了约 75%。这叫过度自信：信心跑得比准确率快，是人脑出厂自带的系统性偏差，在专家身上同样存在（有些研究里更严重）。

经典的及格样本是天气预报员。美国国家气象局几十年前就开始给「降水概率 30%」这类预测做校准统计：所有报 30% 的日子，真的约 30% 下了雨。为什么他们能行？三个条件：**预测被逐条记录、结果及时反馈、同一类事件反复出现**。这恰好是大多数人生决定缺的三样东西——你的「这次创业七成能成」永远没人对账。

Brier 分数：给每个概率打的分

校准只是一半。一个永远说「50%」的人在抛硬币预测上校准完美，但毫无价值——他不提供区分度。所以需要一把同时奖励「准」和「敢」的尺子：Brier 分数，每题得分 = （你报的概率 - 实际结果）²，实际发生记 1、没发生记 0，分数越低越好，平均一大批预测就是你的水平。

平方是精髓。报 90% 而结果没发生，损失 $(0.9)^2 = 0.81$ ；报 60% 而没发生，只损失 0.36。**过度自信被平方项加倍惩罚**，所以 Brier 分数下最优策略永远是报出你的真实信念——想靠「往中间缩」刷分反而更糟（这种性质叫严格适当，行话叫 proper scoring rule）。它把「诚实」变成了数学上的最优解。

预测锦标赛：赢家用的方法

2011 年起，心理学家菲利普·泰特洛克（Philip Tetlock）办了著名的「良好判断计划」，参加美国情报界（IARPA）主办的预测锦标赛：几百个地缘政治问题（「朝鲜明年会进行核试验吗」），数千名志愿者给出概率，用 Brier 分数排名。结果打败了对照组约 30%，前 2% 的「超级预测者」甚至压过了能接触机密情报的职业情报分析师。

赢家们的方法没有魔法，泰特洛克总结下来核心就这么几条：

- **先看基准率（outside view）**：不问「这事多特殊」，先问「这类事 historically 多常发生」。预测一家创业公司成败，先查同类公司五年存活率（约一半），再从 50% 出发调整——而不是从创始人的激情出发。
- **把问题拆碎（费米估算）**：「他当选的概率」拆成党内提名概率 × 大选条件概率，每个子问题的基准率都好查，误差互相抵消。
- **小步高频更新**：新证据来了就调，每次调一点（10% 而不是 50%），既不装死也不翻烧饼。
- **精确到个位**：超级预测者会报 63% 而不是「六成多」。这不是装腔，是逼自己做区分——「差不多」是校准最大的敌人。
- **对账**：记录自己说过的每一个概率，事后算分。这是唯一不可省略的一条。

大白话

大白话：估概率的手艺 = 从基准率出发 + 小步更新 + 逐条对账。它更像会计而不是占卜：赢家不是看得最远的人，是账记得最诚实的人。

一个能今晚就用的练习

从今晚开始，建一个表格：左边写下你对未来一两周可验证之事的概率预测（「周五前 PR 会合并：70%」「这个月会去健身 8 次以上：60%」），右边到点填结果。攒到三五十条，画一张校准图。你会亲眼看到自己的过度自信长什么样——多半「90%」的那栏只中了 70%。然后调。

这个练习改变的不是数字，是关系：你和「不确定」之间，从一笔糊涂账变成了一本能对账的账。下一章讲对账之后怎么办——证据来了，信念具体该怎么动。

第十章 · 做一个会改主意的人

信念不是开关，是旋钮

第三章我们把贝叶斯定理当成一次性的算术：先验 + 证据 = 后验。但真正的用法是把那个公式转成一个循环：今天的后验，是明天的先验。证据一条条来，信念一格格调。这就是贝叶斯更新作为生活方式的含义：**信念不是 0 和 1 之间的开关，而是一个永远停在某个刻度上、随时可被证据拧动的旋钮。**

「改主意」在口语里近乎贬义（立场不坚），在贝叶斯框架里它是唯一理性的动作——证据动了你不动，错的是你。凯恩斯的另一句名言放在这里：「当事实改变，我改变我的想法。你呢，先生？」

先验从哪来：没人从白纸开始

最常见的反对是：「先验太主观了，公说公有理。」这个反对对了一半：先验确实主观，但三条实践规则能把它关进笼子：

- **先用基准率当默认先验**（第九章的 outside view）。「这家餐厅会好吃吗」的先验不是 50%，是「这条街的餐厅里我看走眼的比例」。先验不该是愿望，该是档案。
- **先验弱没关系，证据会赢**。两个起点不同的人，一个先验 20%、一个 80%，只要看同一条够强的证据并诚实更新，后验会收敛到一起。行话叫「先验被证据淹没」。分歧长期存在，多半不是先验不同，而是有一方没在更新。
- **避开 0 和 1**。把先验设成 0%（绝不可能）或 100%（铁板钉钉），数学上没有任何证据能再移动它——贝叶斯公式里， $\text{后验} = \text{先验} \times \text{似然} / \text{归一化常数}$ ，先验为零则永远为零。这就是为什么教条在认识论上是死局：**它不是错，它是不再可改**。给一切留万分之一的活口，是理性的自保。

证据有多强：看「这个世界里多常见」

更新的幅度由似然比决定：这条证据在「假设为真」的世界里出现的概率，除以它在「假设为假」的世界里出现的概率。大白话：**证据的力度 = 它有多「挑世界」。**

用赔率形式看贝叶斯定理，这一点特别干净：

$$\text{后验赔率} = \text{先验赔率} \times \text{似然比}$$

例子：桌上有一枚正常硬币和一枚两面都是头的硬币，随机拿一枚掷出正面。这条证据的似然比 = 1（双面币必出正面） $\div$ 0.5（正常币一半出正面）= 2。先验赔率 1:1，掷出一个正面后后验赔率 2:1，即 67% 是双面币。再连掷三个正面，似然比 $2^3 = 8$ ，赔率变 8:1（89%）。**每乘一次似然比，旋钮转一格**——证据对信念的作用是乘法，不是加法。

反过来，「两个世界里同样常见」的证据似然比是 1，转不动旋钮。「他很焦虑」对「他是不是骗子」几乎没有证据价值——无辜者在审讯室里也焦虑。分辨证据，先问：*如果他是无辜的，我还会看到这一幕吗？*

该更新多少：多，但不翻烧饼

第九章超级预测者的「小步高频」在这里有了数学解释。证据通常是含噪声的：一条新闻可能是误报，一个样本可能是巧合。对含噪声的证据做满额更新，就会被随机游走带着翻烧饼。贝叶斯的做法是把「证据本身的可靠度」也折进似然比——可靠度打七折，更新幅度就打七折。**更新的幅度应该正比于证据的区分度乘可信度**，而不是正比于新闻的标题字号。

什么时候该「固执」

贝叶斯主义者常被嘲讽为墙头草。其实框架自己说出了固执的合法条件：

- **先验极强、证据极弱时，不动是对的。**「有人声称永动机造出来了」——先验（热力学定律两百年的全胜纪录）太强，单条声称的似然比推不动它。非凡的主张需要非凡的证据，这不是保守，是算术。
- **证据被污染时，不动是对的。**审讯里「嫌疑人在压力下供述」这条证据，在无辜世界里也会以不小概率出现，似然比接近 1，旋钮不动。

但注意这两种固执都给出了**明确的解锁条件**：更强的证据、更干净的来源。教条的固执没有解锁条件——这就是科学信念与意识形态的分界线：前者知道自己在什么情况下会改，后者不知道。

大白话

大白话：贝叶斯式的人生：给一切信念标一个 0 到 1 之间（不含两端）的数；每来一条证据，问它在两个世界里各有多常见，按比例拧旋钮；永远答得出「什么证据会让我改主意」。答不出这个问题的信念，不是信念，是纹身。

到这里，「怎么想」的部分齐了。剩下两章进入「怎么做」：知道概率之后，怎么下注、怎么活。

第十一章 · 期望值是起点，不是终点

期望值不够用的时候

第五章留下的账现在来收。期望值的决策规则是「选期望最大的」，它在两个地方会骗人：一是它把分布揉成一个数（第五章）；二是——这一章的主角——**真人面对钱的时候，效用函数长得和教科书假设不一样，而且错得有规律。**

1979 年，卡尼曼和特沃斯基发表前景理论（prospect theory），用一系列精巧的实验画出了真人决策的「实测曲线」。它配得上概率书里的一章，因为它是「人类实际如何用概率做决定」最经得起复现的描述。核心就四根钉子。

钉子一：参照点——没有绝对的贫富，只有相对的输赢

你评价的不是「我最终有多少钱」，而是「我比参照点多还是少」。年薪涨了两万但同事涨了三万，你感到的是**损失**。同一个结果，参照点挪一下，体验从盈变亏。期望效用理论里没有参照点这个位置——这是它和真人的第一道裂缝。

钉子二：损失厌恶——丢一百的痛约等于捡两百的爽

实验反复测出：损失带来的负效用约为等量收益正效用的 2 倍左右。这直接解释了为什么「50% 赢 200 / 50% 输 100」这种正期望赌局多数人不肯玩——不是不懂期望，是输 100 的痛在效用账上约等于 -200，整个赌局在**效用期望**上是负的。

损失厌恶本身未必是「错」——第六章的乘法世界给了它一个理性辩护：深亏会削掉复利基部。真正的问题在下一根钉子。

钉子三：参照点两侧，风险偏好反转

实验发现人在**收益区厌恶风险**（宁要确定拿 500，不要 50% 拿 1000），在**损失区追逐风险**（宁要 50% 亏 1000，不愿确定亏 500）。合起来是一条 S 形价值曲线：收益区凹、损失区凸，损失侧更陡。

这个反转在现实里杀伤力极大。它解释了交易大厅最古老的经验——「人总是过早止盈、死扛亏损」：浮盈时你在收益区，厌恶风险，急着落袋；浮亏时你在损失区，追逐风险，宁可赌它涨回来也不认亏。Disposition effect（处置效应）在无数市场的真实交易数据里被反复验证。**这不是意志薄弱，是 S 形曲线的机械后果。**

钉子四：概率加权——人对概率本身也是扭曲的

前景理论的第二根轴：人不按真实概率加权，而按一条扭曲的概率加权函数——**高估小概率，低估中大概率，对「确定」和「不可能」两个端点格外敏感。**

- 高估小概率 + 收益区 → 买彩票：百万分之一的大奖在心理上被放大成「万一呢」。
- 高估小概率 + 损失区 → 买保险：百万分之一的灾难被放大成「睡不着」。
- 第五章那个「人为什么同时买彩票和保险」的谜题，到这里被一根概率加权曲线同时解开。

端点敏感也有名有姓：**确定性效应**——从 99% 提到 100% 的心理跳跃，远大于从 40% 提到 41%。「保证」两个字之所以值那么多钱，是因为它在加权函数上跳崖。

那该怎么办：用系统覆盖出厂设置

知道了 S 形曲线，不是让你「凭意志克服」——没有人能凭意志克服自己的出厂设置。能做的是用规则和流程把决策外包给清醒时的自己：

- **预设止损/止盈，写下来再入场**：把「什么时候卖」从损失区的自己手里没收。
- **换参照点**：以「全部身家」而不是「这一单」为参照系评估盈亏——单笔亏损在总资产框架里会缩小到理性尺寸。
- **把概率写出来再决定**：第九章的练习在这里回报——白纸黑字的 63% 能顶住「万一呢」对小概率的放大。
- **问期望，也问分布**：期望值是起点；第五章的方差、左尾、偏度三个追问补全后半程。

大白话

大白话：前景理论不是教你「别犯错」，是给你一张自己会在哪犯错的地图。地图救不了不看的司机，但能让看的人提前装护栏。

还剩最后一个问题，也是全书最硬的一个：下注多少。就算概率估对了、效用也想清了，押多大的注，仍能把正期望的好局玩成归零。最后一章，凯利判据。

第十二章 · 先别死，再谈赢

最后一问：押多少

全书至此，概率会定义了、会算了、会估了、会更新了，决策的坑也标出来了。但还差临门一脚：一个赌局摆在你面前，概率和赔率都清楚——**你该押上身家的百分之几？**

这个问题比看起来生死攸关。回到第六章那个乘法赌局：赢了 $\times 1.5$ ，输了 $\times 0.6$ ，期望 $+5\%$ 。现在知道它是正期望的好局。全押会怎样？我们已经算过：五赢五输后剩 0.59 ，长期趋向归零。**正期望 + 全押 = 数学保证的破产**。问题出在没有选择仓位。

凯利判据：最大化时间平均的增长率

1956 年，贝尔实验室的约翰·凯利（John Kelly）在研究通信信道噪声时顺手解决了一个赌徒问题：如果每次都能押上资产的某个比例 f ，长期来看哪个比例让资产增长最快？答案是对数效用的解——最大化**每局对数收益的期望**。对「赢概率 p 、赔率 b （赢 1 赔 b ）」的简单赌局：

$$f^* = p - (1-p)/b$$

这就是**凯利判据**。读一个数：胜率 60% 、赔率 1 （输赢等额）的赌局， $f^* = 0.6 - 0.4/1 = 0.2$ ——每局押身家的 20% 。不是 100% （那是破产特快），不是「越多越好」，是一个精确的、由概率和赔率决定的中间值。

为什么对数？第六章给了答案：乘法世界里，个体的命运是时间平均，而 $\ln(\text{资产})$ 的期望增长正是时间平均增长率的度量。凯利判据不是「一个还不错的启发式」，它是遍历性数学的必然结论：**在乘法世界里活得最好的人，就是对数效用**人——不管他嘴上信什么效用理论。

凯利的两张面孔

面孔一：它比其他任何比例长得都快。长期看，凯利仓位的资产曲线几乎必然超过任何固定比例的替代方案。索普（Ed Thorp）——《战胜庄家》作者、后来最成功的对冲基金经理之一——把它从二十一点牌桌搬进金融市场，三十年无亏损年度的纪录是这套数学最硬的广告。

面孔二：它的颠簸超出多数人的承受。全凯利仓位的回撤又深又久——账面腰斩是常态。更危险的是，凯利公式对你的**概率估计极其敏感**：你以为胜率 60% 实际只有 55%，你以为在按 20% 下注，其实已经是「过度下注」——过度下注比下注不足惩罚重得多（增长率跌得更快，且可能转负）。实战中的通行做法是**半凯利**甚至四分之一凯利：增长率为全凯利的 3/4，波动只有一半。在没有把握的估计面前主动减速，是凯利系统自带的谦卑。

毁灭：一条不可逾越的红线

凯利框架的深处是全书最后一个、也是最重要的概念：毁灭（ruin）——资产归零（或跌破不可恢复线）的状态。它的特殊性在于**不可逆**：亏损 50% 可以赚回来，归零之后没有「之后」——你出局了，遍历性对你失效，大数定律对你失效，一切「长期」承诺都作废。

塔勒布把这条讲成了一条决策总则：**任何带有毁灭分支的选项，无论期望值多高，都不进入候选集**。期望值是「平行宇宙的平均」，而毁灭意味着「在你这个宇宙里游戏结束」。俄罗斯轮盘赌一局 100 万美元——期望正得诱人，但你应该拒绝，不是因为它期望不好，而是因为它的左尾有一条不归路。先别死，再谈赢；顺序不能反。

大白话

大白话：把十二收成一句操作纪律：① 期望值负的局不玩；② 期望正但有毁灭分支的局不玩；③ 剩下的局，按凯利的一小部分下注，且宁少勿多；④ 让时间和复利替你工作——你唯一要做的就是保证自己始终在场。

尾声：与不确定共处

回到导读里那个问题：「明天降水概率 70%」到底是什么意思？现在你可以回答：它是对一个单次事件量化的信念，用法拉普拉斯数不了格子、频率派数不了重复，但它照样能进柯尔

莫哥洛夫的公理、能用 Brier 分数对账、能进你的决策。

而那个贯穿全书的大问题——未来不确定时理性的人怎么活——十二道答案合起来是这样：把信念量化成概率并诚实对账（九、十），用价格和市场校准自己（七），尊重肥尾、对「百年一遇」保持警觉（八），用期望值做初筛但不迷信它（五、十一），永远问分布的形状、问左尾有没有不归路（十二），在乘法的世界里按时间平均而非系综平均思考（六），用凯利的一小部分决定押多少（十二）。

概率不是消除不确定的法术。它是让你在不确定中仍然能行动、能改错、能活下来的记账系统。世界不会因此变得可预测——但你会变得难被击倒。这已经是一个人能向随机性要到的最好条件了。

概率：与不确定共处 · 王建硕