

跌多深、多久、为什么

标普 500 历史回撤、恢复时间与度量口径

Vincent Wang

2026 年 7 月

研究的出发点

98.5 年日度数据、73 次水下事件：
在下一次回撤到来之前，历史能告诉我们什么？

声明

本材料仅用于研究，不构成投资建议。

全部结果为历史数据的描述性统计，历史回撤不预示未来结果。

| 长期投资者的回撤四问

01 | 深度



多深

多数回撤停在 5-10%，但少数系统性熊市决定了风险预算的上限。

02 | 时长



多久

恢复期以年计、很难说准；不过大半跌幅往往在前半程就已收复。

03 | 归因



为什么

有没有赶上经济衰退，决定了这次回撤有多深、拖多久。

还有贯穿全文的第四问：用什么口径度量。同一段历史，换个口径，答案就会不同。

| 数据与指数沿革

四层数据

- ▶ 日线核心：Yahoo ^GSPC 收盘价，1927-12-30 至 2026-07-02
- ▶ 月线：Shiller 综合序列（1871 年起，月内均值）
- ▶ 口径变体：名义/实际 \times 价格/全收益
- ▶ 宏观与跨资产：NBER 衰退月份、九只 ETF

24,742
个交易日

98.5 年
日度样本跨度

73 次
5%+ 水下事件（ATH 锚定）

! 1957 年前日线为前身指数 S&P 90 的回溯数据；1926 年前月线来自 Cowles 重构。

全部输入均通过自动数据校验（日期、缺口、单日极端涨跌幅）。

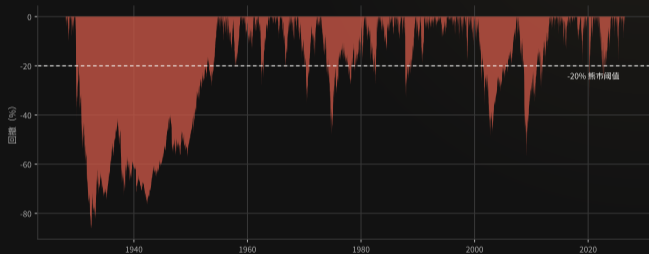
| 事件定义：从前高到收复前高

一次水下事件

$$D_t = P_t / M_t - 1$$

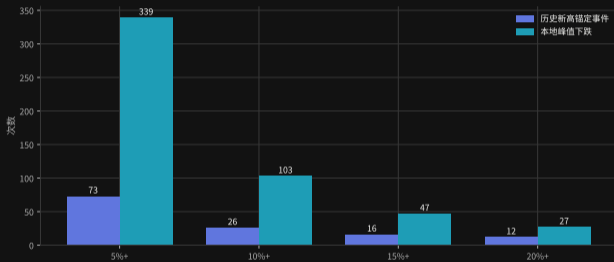
从前高跌落 → 谷底 → 重新收复前高；按最大跌幅归入 5/10/15/20% 档，一轮熊市只计一次。

全部用日收盘价；时间按日历天计数；尚未收复的事件按截尾处理（当前为 0 个）。



怎么定义就怎么计数：后面所有数字都建立在这个定义上。

同一份数据，两种答案



73 vs 339

5%+: ATH 锚定 vs 本地峰值

3.4 次/年

本地峰值口径下的 5% 回调频率

1937-38 (-54.5%)、2011 (-19.4%)

这类著名下跌，都被并入了更长的水下事件，不再单独计数。

本文的事件计数是下限口径，别拿去和“每年三次 5% 回调”这类说法直接对比。

| 执行摘要：四个数字

73 / 26 / 12

5%+ / 10%+ / 20%+ 事件数 (ATH 锚定)

**-35.0% vs
-14.3%**

深度中位数：衰退重叠 vs 非重叠

764 天

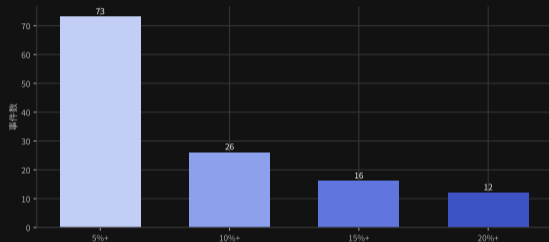
20%+ 恢复中位数；95% CI [666, 2310] 天

5.4 → 2.7 年

计入股息后，20%+ 熊市水下时间中位数

跌多深、多久、为什么、用什么口径——四个问题要一起回答。

| 阈值频率与抽样不确定性



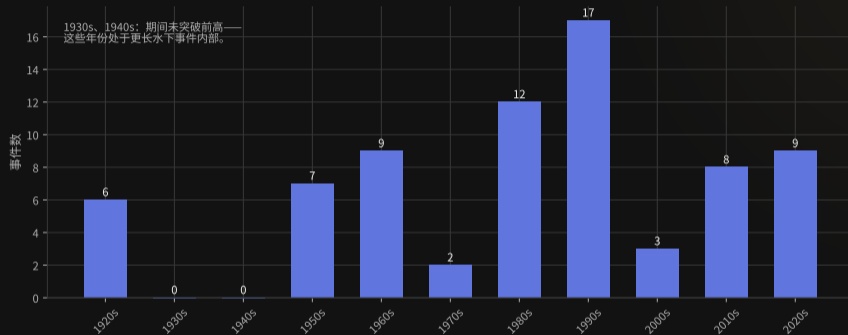
12 个
20%+ 档的全部样本

[666, 2310] 天
恢复中位数的 bootstrap 95% CI

10,000 次重抽样、固定种子；区间上下限差了三倍多。

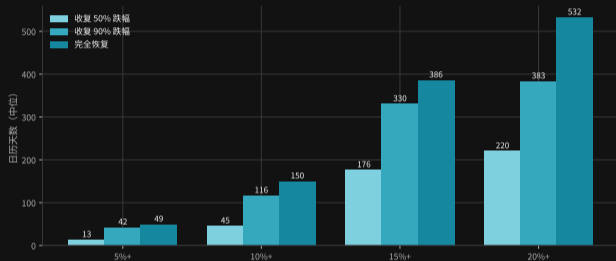
做规划用区间，别只用一个中位数。

| 分年代分布：扎堆出现，而非均匀分布



1930s, 1940s 的零事件是口径造成的：那些年份整段泡在更长的水下事件里，并不是市场平静。

恢复里程碑：大部分跌幅在前半程就已收复



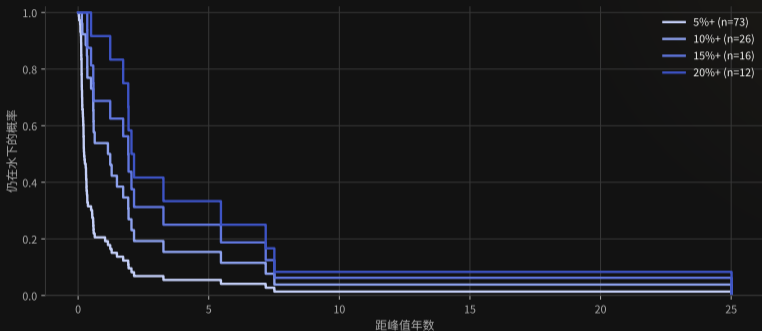
220 天
收复一半跌幅 (中位)

383 天
收复九成跌幅

532 天
完全收复前高

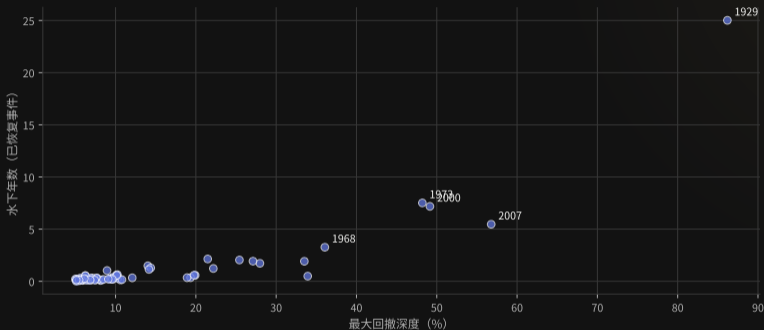
取钱、再融资、对外沟通，可以按“收复一半/九成”的节点安排，不必都等到重回前高那一天。

| 生存视角：跌下去之后，多久能爬出水面？



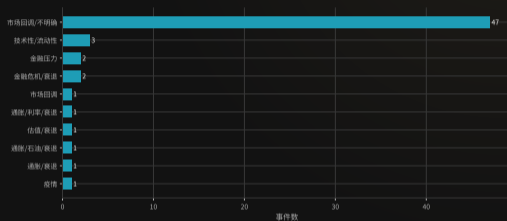
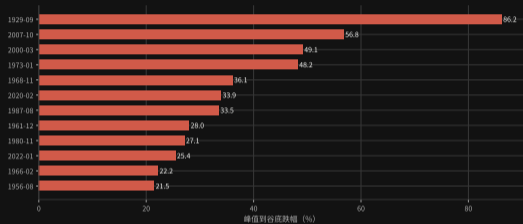
Kaplan-Meier 估计，自动处理截尾事件。20%+ 组的长尾来自 1929、1973、2000、2007 四次系统性熊市。

| 深度与恢复时间：相关，但绝不是一回事



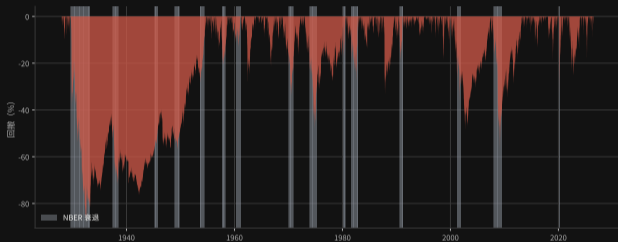
跌幅相近的两次危机，恢复时间可以差好几倍（比较 2020 和 1973）——深度和时长要分开看、都要报。

| 最深熊市与原因结构



标签覆盖全部 26 个 10%+ 事件（人工整理、带置信度，不是因果检验）。深的回撤多与金融危机和衰退相伴；纯技术性/流动性事件（1987、2018-02）跌得深但恢复快。

宏观环境：有没有衰退，是深度的分水岭



衰退重叠 (10 次)

-35.0% | 988 天

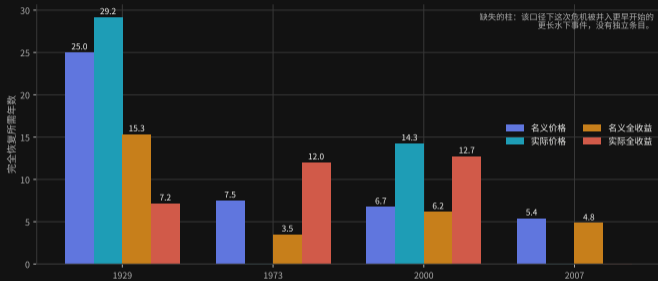
无衰退重叠 (16 次)

-14.3% | 216 天

重叠只是相关性描述，不构成因果识别。

宏观判断先问“会不会衰退”，再问“会不会回撤”。

| 换一种度量口径，恢复叙事完全不同



25.0 → 7.2
年

1929：名义价格 → 实际全收益

7.5 → 12.0
年

1973：算上通胀，实际恢复更久

实际口径下 2000-2013 是一段连续水下期，两次崩盘都在其中——用什么口径汇报，要和资金的实际用途一致。

跨资产传导：危机中的相对表现

标普500 (价格指数)	-12	-49	-57	-14	-10	-20	-34	-25	-19
SPY (标普500)	-12	-48	-55	-13	-10	-19	-34	-24	-19
QQQ (纳斯达克100)	-12	-83	-53	-16	-11	-23	-29	-35	-23
IWM (罗素2000)	无数据	无数据	-58	-26	-9	-26	-41	-27	-23
XLK (科技)	-12	-82	-53	-14	-10	-24	-31	-33	-26
XLF (金融)	-21	-37	-82	-22	-12	-22	-43	-26	-16
XLE (能源)	-15	-42	-57	-34	-14	-30	-56	-26	-19
XLU (公用事业)	-10	-52	-46	-11	-7	-9	-36	-25	-8
TLT (20年期以上美债)	无数据	无数据	-27	-6	-5	-4	-16	-40	-9
GLD (黄金)	无数据	无数据	-29	-13	-13	-5	-13	-21	-7
	1999-07	2000-03	2007-10	2015-05	2018-01	2018-09	2020-02	2022-01	2025-02

！ 国债的对冲效果取决于危机类型：2022年通胀型熊市里，TLT跟着股票一起跌了40%。图中为调整后收盘价（含分红）；“无数据”表示该资产当时尚未成立。

| 分层风控：不同深度是不同的问题

5% 回调 = 噪音

本地峰值口径下平均每年约 3.4 次。

不需要任何动作。

10%+ = 触发纪律

按既定预案沟通、再平衡、控制杠杆。

会不会继续变深，主要看衰退判断。

20%+ = 资本保护

恢复中位 764 天，置信区间以年计。

用区间和里程碑做规划，别押单点。

把这三档分别写进风控手册：什么情况触发、做什么、怎么对外说。

| 给防御策略当对照组

本研究不是择时信号

它是所有防御方案的对照组：趋势过滤、波动率控制、期权保护，都应该说清楚自己改善了什么——深度、水下时间还是里程碑——以及改善了多少。

下一步

- ▶ 利率/通胀环境下的正式条件统计
- ▶ 给 ETF 面板补上恢复时间维度

没有对照组，防御策略就无从评估。

| 局限性

！核心日度序列是价格指数；全收益/实际值稳健性用的是月度数据（月内均值会抹平极端）。

！原因标签是叙事性归因；衰退重叠是相关性，不是因果。

！ETF 面板用调整后收盘价，各资产成立时间不一；1957 年前日线为 S&P 90 回溯、1926 年前月线为 Cowles 重构。

完整的局限清单与可复现性说明见论文第 12 节。

长期收益，来自承担并穿越一次次回 撤。

历史给不了预测，但可以校准预期：
多深·多久·为什么·用什么口径

Vincent Wang | vincentwang.research@outlook.com