資金流與「跟大戶」：為何高命中率不等於正期望值

賽馬彩池的數學檢驗｜Sigma Quant 研究報告｜2026 年 9 月

摘要

本文檢驗一個常見說法：只要讀得懂資金流、跟隨大戶，就能在賽馬中長期獲利。我們刻意採用三個偏強的假設——大戶必知賽果、大戶準確知道機率、大戶只知道近似機率——再逐一證明：跟隨者的期望回報不會高於零；當資金訊號帶有雜訊時，更會低於全場平均的 −τ。合成數據模擬（10 萬場、每場 12 匹）顯示，「跟資金最重的一匹」勝出率約 25.5%，遠高於隨機的 8.3%，但平均回報約 −22.6%。資金流的合理用法，是與自己的機率估計按精度加權融合，而不是取代它。

> **三個結論**
> 1. 命中率高，不等於值博。
> 2. 資金流是資訊，不是價格優勢：你看到訊號時，價格通常已經反映。
> 3. 有自己的機率模型時，用資金流校正它；沒有時，不要單靠資金流落注。

1 設定與符號

彩池投注中，馬匹 i 的勝出概率為 p_i；該馬的勝彩池資金為 W_i，總彩池為 P，資金佔比 $s_i = W_i/P$。抽水率記為 τ（本文以 17.5% 作示範參數，實際數字請以香港賽馬會最新投注規則為準）。派彩由扣除抽水後的彩池按中獎資金分配，並非事前固定。為聚焦核心，我們忽略派彩取整、多匹馬同時衝線（並列）及自己下注對賠率的影響。

$$D_i = \frac{1-\tau}{s_i}, \qquad s_i = \frac{W_i}{P} \tag{1}$$

每 $1 本金的回收 R_i 只在該馬勝出時等於 D_i，否則為零。因此：

$$E[R_i] = p_i D_i = \frac{(1-\tau)\,p_i}{s_i} \tag{2}$$

定理 1（恆等式）：全場按資金加權的平均回收恆為 $1-\tau$。

$$\sum_i s_i E[R_i] = (1-\tau)\sum_i p_i = 1-\tau \tag{3}$$

證明：對每匹馬取式 (2) 乘以 s_i，s_i 互相抵消，剩下 $(1-\tau)\Sigma p_i$；而全部馬匹的勝出概率總和為 1。■

換言之，整體跟著資金走，平均必然損失 τ。只有個別馬匹被市場低估（資金佔比低於下式門檻）時才有正期望值：

$$E[R_i] \geq 1 \iff s_i \leq (1-\tau)\,p_i \tag{4}$$

例 1：四匹馬的迷你彩池

馬匹	資金（萬元）	佔比 s	派彩 D	真實 p	期望淨回報
A	40	40%	2.06	30%	−38.1%
B	30	30%	2.75	35%	−3.8%
C	20	20%	4.13	22%	−9.3%
D	10	10%	8.25	13%	+7.3%
加權平均	100	100%	—	100%	−17.5%

資金最重的 A 期望回報最差：其佔比 40% 遠高於門檻 $(1-\tau)p = 24.75\%$。只有資金最輕的 D 低於門檻。加權平均正好等於 $-\tau$，與定理 1 一致。

© Sigma Quant 數馬講　版權所有，未經許可不得轉載或改作｜sigmaquanthk.com

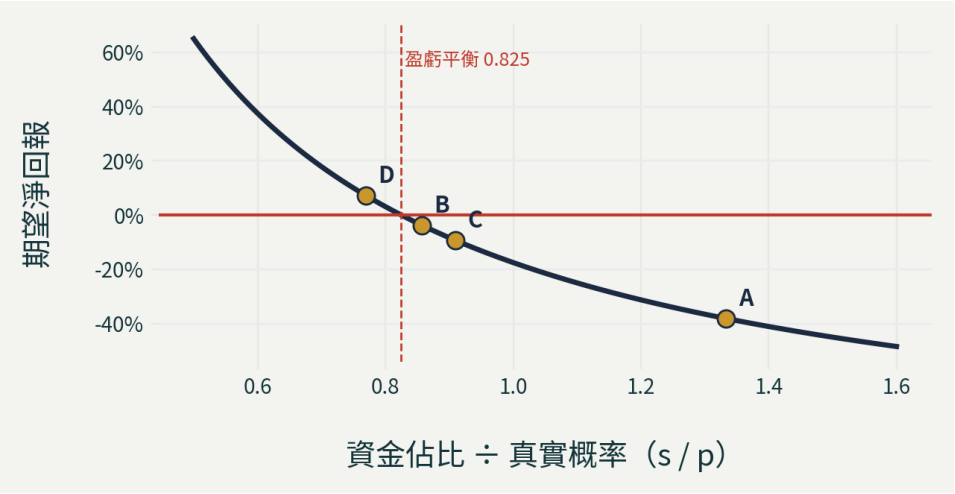


圖 1 期望淨回報只取決於「資金佔比 ÷ 真實概率」。四匹馬取自例 1；紅線 = 盈虧平衡（比值 0.825）。

2 三個大膽假設，逐一檢驗

假設 A：大戶必知賽果，且資金無限、互相競爭。知道贏家的人，只要 $s_w \le 1-\tau$ 就有利可圖，競爭會把該馬資金佔比推到 $1-\tau$，派彩趨近 1.00：

$$D_w = \frac{1-\tau}{s_w} \ge 1 \iff s_w \le 1-\tau \Rightarrow D_w \to 1.00 \tag{5}$$

實際上，大部分勝出馬的派彩明顯高於 1.00，所以假設 A 不成立。即使成立，跟隨者最多也只是不賺不賠。

假設 B：大戶準確知道每匹馬的真實概率 p。設某馬原有非知情資金 W，總彩池 P，大戶追加 x。由於 W < P，賠率隨 x 嚴格下降；競爭使大戶加碼至期望值歸零：

$$o(x) = \frac{(1-\tau)(P+x)}{W+x}, \qquad o'(x) = \frac{(1-\tau)(W-P)}{(W+x)^2} < 0 \tag{6}$$

$$x^* = \frac{p(1-\tau)P - W}{1-p(1-\tau)} \Rightarrow p \cdot o(x^*) = 1 \tag{7}$$

證明：式 (6) 的導數恆為負，故 $p \cdot o(x) - 1$ 隨 x 單調遞減；令其為零即得式 (7)。任何在 x^* 之後才入場的跟注者，期望回報必為負。■

例 2：一匹 50% 勝率的馬

設 p = 50%，P = 100 萬，非知情資金 W = 35 萬，τ = 17.5%。起初賠率 2.36，期望淨回報 +17.9%。大戶加碼至 $x^* \approx$ 10.6 萬時，賠率跌至 2.00，期望值歸零；此時你再跟注 1 萬，賠率變 1.97，期望淨回報 −1.3%；若跟注 5 萬，賠率 1.88，期望淨回報 −5.8%。大戶就算完全正確，你買入時的價格已不再值博。

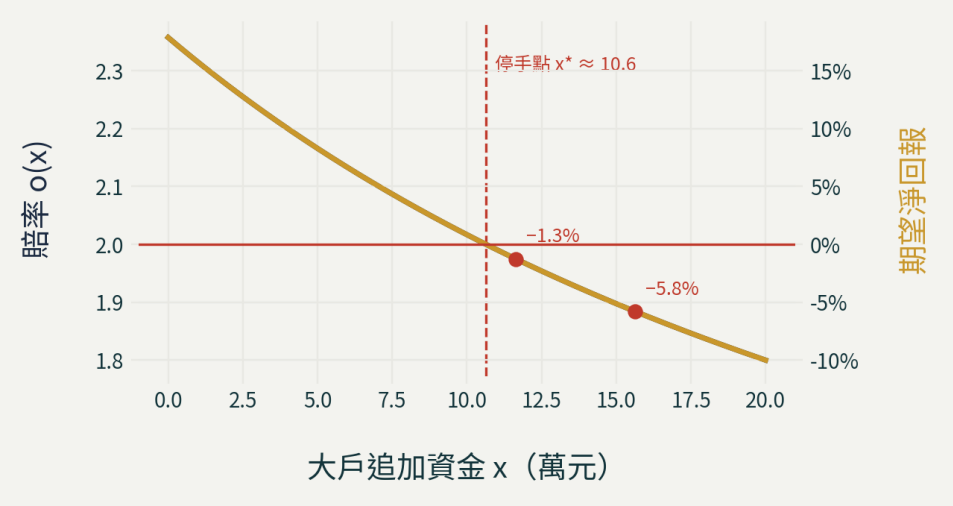


圖 2 例 2 的賠率與期望淨回報路徑。虛線 = 大戶停手點 x*；其後入場者期望值為負。

假設 C：大戶只知道近似機率。把資金佔比視為「真實概率＋雜訊」，並用對數常態近似：

$$\ln s_i = \ln p_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma_\varepsilon^2), \quad \ln p_i \sim N(m, \sigma_u^2) \tag{8}$$

定理 2（資金越重，回報越被高估）：

$$E[p \mid s] \propto s^{\kappa}, \quad \kappa = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\varepsilon^2} \Rightarrow E[R \mid s] = C s^{\kappa - 1} \tag{9}$$

證明：在高斯模型中，$E[\ln p \mid \ln s] = m + \kappa(\ln s - m)$，其中 $\kappa < 1$，即條件期望向平均值收縮；取指數後 $E[p \mid s] \propto s^{\kappa}$，再除以 s 得式 (9)。■ 直觀地說：資金佔比最高的馬，有一部分只是「正向雜訊」的結果，跟注等於承受贏家詛咒。

例 3：資金佔比加倍的代價

雜訊比重 1 − κ	佔比 ×2 時期望回收係數	佔比 ×5 時期望回收係數
0.07	0.953	0.893
0.15	0.901	0.786
0.30	0.812	0.617

例如 $1 - \kappa = 0.15$ 時，資金佔比由 10% 升至 20%，期望回收乘以 0.901，即再少約 9.9%；升至 50% 則乘以 0.786。這個係數還未計入抽水。

3 模擬與實證

以下為**合成數據模擬**，並非香港賽馬會歷史數據，目的是驗證機制。設定：10 萬場、每場 12 匹；ln 實力 ~ $N(0, 0.9^2)$ 後歸一化為 p；資金佔比 $\ln s = \ln p + \varepsilon$，$\sigma_\varepsilon = 0.25$；自己的模型 $\ln p_m = \ln p + \delta$，$\sigma_\delta = 0.35$；$\tau = 17.5\%$；每注固定 $1。

策略	每注命中率	平均回報 ± 標準誤
只買資金最重的一匹	25.5%	−22.6% ± 0.5%
買資金最重的三匹	17.8%	−20.4% ± 0.3%
自己模型第一選	24.6%	−13.7% ± 0.6%
└ 與資金第一相同（62.6% 場次）	28.8%	−19.9% ± 0.5%
└ 與資金第一衝突（37.4% 場次）	17.6%	−3.4% ± 1.2%
自己模型，優勢 > 5% 才買	7.7%	+2.5% ± 0.9%
融合（w = 0.34），優勢 > 5% 才買	6.6%	+14.7% ± 2.4%

隨機選馬的命中率是 1/12 ≈ 8.3%，所以跟資金確實「勝率高」；但回報 −22.6%，比全場平均 −17.5% 更差，正如定理 2 所預測。我們再用二十等分位檢驗指數：擬合 ln(勝出率 / s) 對 ln s 的斜率為 −0.0713，而理論值 $\kappa - 1 = -0.0716$。

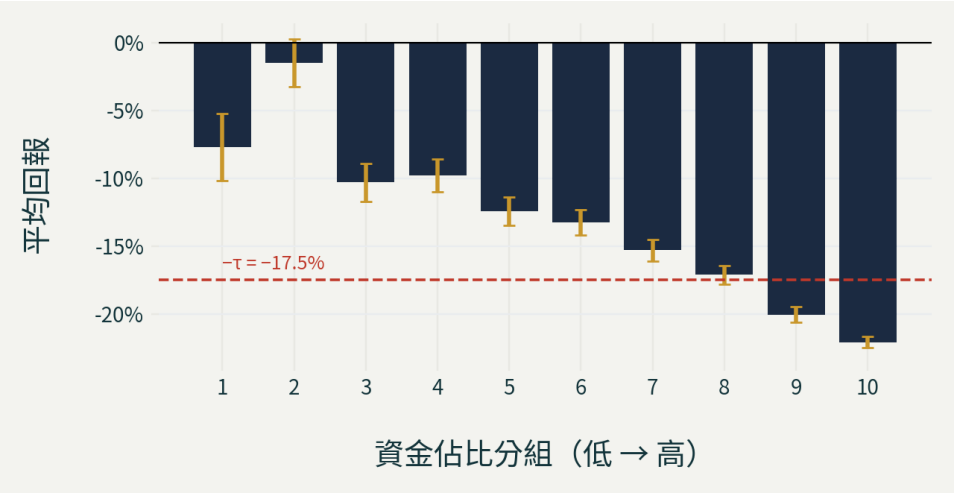


圖 3 按資金佔比分十組的模擬平均回報。整體而言資金越重，回報越低（低佔比組別誤差較大）；全部低於盈虧平衡，虛線 = −τ。

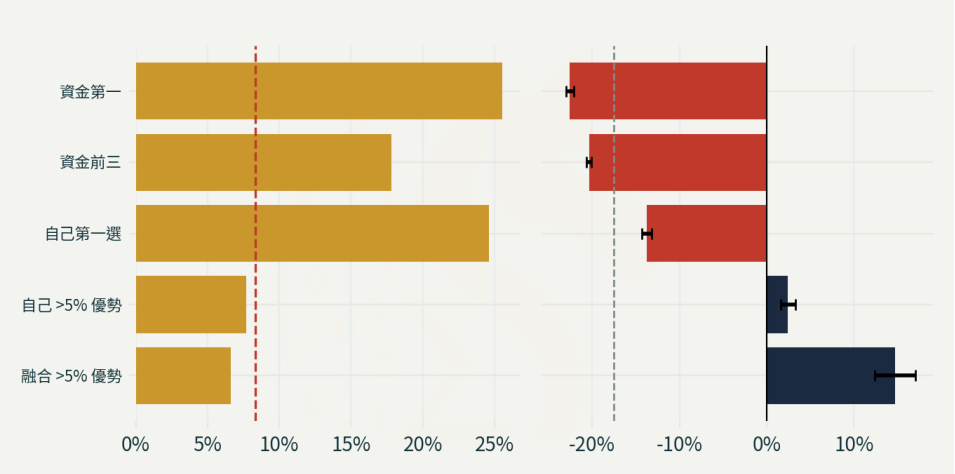


圖 4 各策略的每注命中率與平均回報（誤差線 = ±1 標準誤）。高命中率與正回報可以毫無關係。

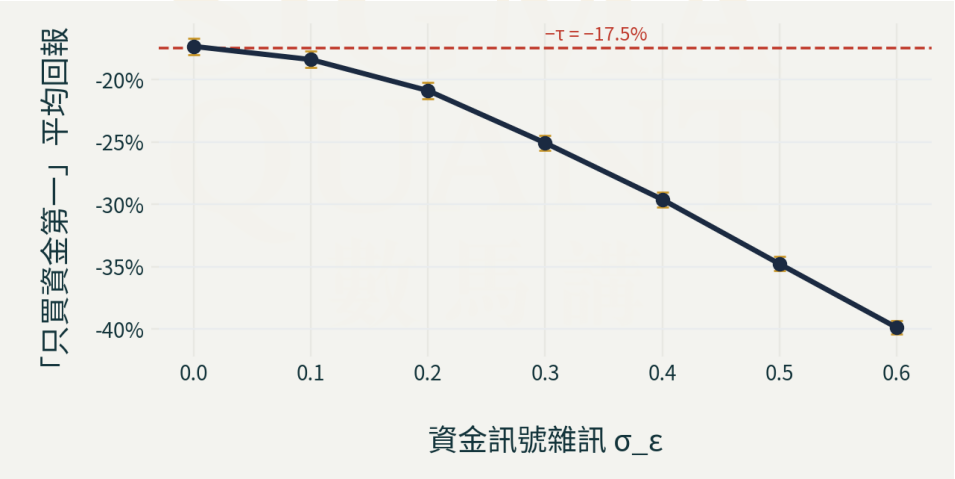


圖 5 資金訊號雜訊 σ_ε 愈大，「只買資金第一」的平均回報愈差；σ_ε = 0 時回到 −τ。

文獻方面，投注市場中「熱門—冷門偏差」（冷門的期望回報較差）是被廣泛記錄的現象，Ali (1977) 已從分歧信念解釋其成因。Sigma Quant 較早的研究亦指出：市場價格水平仍是主訊號，賠率路徑只帶來小幅增量，尚未證明投注優勢。

4 資金流的正確用法

若自己有機率估計 p_m（誤差 σ_δ），資金佔比 s 是另一個獨立誤差來源（σ_ε），最佳做法是按精度加權融合，並要求優勢超過安全邊際 m：

$$\ln p_b = w \ln p_m + (1-w)\ln s, \quad w = \frac{\sigma_\varepsilon^2}{\sigma_\delta^2 + \sigma_\varepsilon^2}, \quad \mathrm{Var} = \frac{\sigma_\delta^2 \sigma_\varepsilon^2}{\sigma_\delta^2 + \sigma_\varepsilon^2} \tag{10}$$

$$\text{Bet if } \frac{(1-\tau)\, p_b}{s} > 1 + m, \qquad f^* = \frac{p_b\, o - 1}{o - 1} \tag{11}$$

例 4：模型與資金衝突時

情境	自己 p_m	資金 s	融合 p_b	只信自己	融合後
模型看好，資金冷淡	20%	10%	12.7%	+65.0%	+4.4%
兩者接近	12%	10%	10.6%	−1.0%	−12.2%
模型看淡，資金熱	8%	14%	11.6%	−52.9%	−31.8%

第一行：只信自己會覺得有 +65% 的優勢，凱利注碼約 8.97%；融合後優勢降至 +4.4%，注碼約 0.61%，縮小約 15 倍。這正是「資金與預測衝突時，寧願跟資金」與「完全忽略資金」之間的理性折衷：以精度為權重，而不是二選一。（權重 w = 0.34 由 $\sigma_\delta = 0.35$、$\sigma_\varepsilon = 0.25$ 計得；表內數字忽略歸一化。）

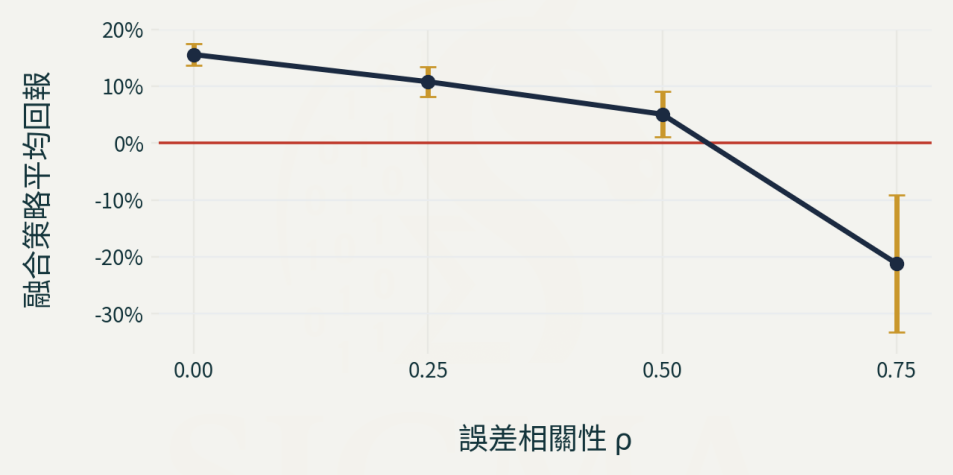


圖 6 融合策略的模擬回報 vs 兩個估計的誤差相關性 ρ。兩者愈相似，融合帶來的額外資訊愈少；ρ = 0.75 時樣本僅約 1,800 注，誤差極大。

情境	建議做法	原因
沒有自己的機率估計	不要單靠資金落注	定理 1、2
資金與模型一致	仍要檢查價格是否符合式 (11)	一致多半已反映在賠率
資金與模型衝突	用融合機率重算，減注或放棄	兩個估計都有雜訊
資金含模型沒有的資訊	作補充，提高權重	精度加權

5 限制與驗證方案

第一，模擬是合成數據，只證明機制，不代表實際回報。第二，假設對數常態與獨立雜訊；若資金與模型的誤差高度相關，融合優勢會消失（圖 6）。第三，忽略自己下注影響、派彩取整、多匹馬同時衝線及不同彩池的抽水差異。

建議以真實數據驗證：(1) 取 T−2 前後的勝彩池佔比 s，迴歸 ln(勝出率/s) 對 ln s，檢驗斜率是否為負；(2) 加入 T−5 至 T−2 的資金變動，按賽季分塊交叉驗證，比較 log loss；(3) 報告 ROI 的信賴區間，而非單一數字。

參考：Ali, M. (1977), Journal of Political Economy；Whelan, K., Economics of Betting Markets 講義；香港賽馬會投注規則與投注指南；Sigma Quant Research 01–04（2026）。

本文只作數據分析與教育用途，不構成投注建議，亦不保證任何回報。賽馬活動涉及風險。