

换模型盲测：Opus 5 vs Fable 5.1

对象：批次 4（英冠 8 场，2026-09-01 已完赛）+ 批次 5（西甲 / 法甲 2 场，2026-09-03 开赛）• 复核日期 2026-09-02 • 输入：同一套盘口截图

结论：10 场里胜负只有 1 场翻转（伯明翰 vs 南安普敦），翻转原因是原分析在均势盘上对已定价信息做了修正；首选比分有 6 场不同，其中 3 场原报的比分连"预测方向内"的最可能比分都不是。大小球与双方进球没有一场方向相反。今晚两场全部一致。

这次复核的价值不在"谁更准"（样本太小），而在暴露了两个可机械化拦截的方法差：① 均势盘上一次"合规"的 1.5pp 修正就足以翻转结论；② 报"方向内比分"与报"网格众数"是两套口径，原分析混用了。两者都已写进发布闸门脚本。

先说清楚这次"盲测"的边界

- 批次 4 做不到真正不知道赛果——回填赛果的正是本 session。所以"盲"是程序性的：全部结论只由「截图去水 → λ 反推 → 泊松网格」机械产生，再按方法论硬规则处理（最高项 <45% 禁修正、45–54% 档净修正 ≤1.5pp），**不做任何定性修正、不搜网**（搜出来的全是赛后报道）。这套流程无法被"知道结果"带偏，但证据面比原分析窄——只有盘口 App 内嵌的 H2H 与近况。
- 批次 5 是真盲：截图 + 赛前新闻检索，正常跑六维。
- 两批用的是同一份截图（NAS 场次文件夹 014–023），庄家均为 m617 系盘口，口径可比。

机械管线（每场完全相同）

- 1X2 去水

三项 1/赔率 按比例归一；抽水 = 加总 – 100%。
- λ 反推

用「大 2.5」去水概率二分搜索泊松 λ（期望总进球）。
- λ 拆分

在 λ 固定下搜索主客拆分，使泊松网格的主胜 / 客胜概率与去水 1X2 误差最小。
- 输出

outcome = 1X2 最高项；score = 网格众数；OU / BTTS 概率落在 45–55% 一律标「无倾向」。

脚本：blind.py（本次临时工具）；其中网格众数排序检查已并入发布闸门 check_consistency.py。

批次 4（英冠 8 场）：原预测 vs 盲测

场次	原（Opus 5）	盲测（Fable 5.1）	胜负	首选比分
林肯城 vs 布莱克本	主胜 2-1 [41.5/26.7/31.8]	主胜 1-1 [41.0/27.1/31.8]	一致	不一致
朴茨茅斯 vs 德比郡	主胜 1-1 [46.3/27.7/26.0]	主胜 1-1 [46.9/26.5/26.5]	一致	一致
普雷斯顿 vs 布里斯托城	客胜 1-2 [33.7/26.4/39.9]	客胜 1-1 [34.6/26.4/39.0]	一致	不一致
谢菲尔德联 vs 博尔顿	主胜 2-1 [56.1/25.7/18.2]	主胜 1-0 / 1-1 并列 [56.1/23.7/20.1]	一致	不一致（弱）
斯旺西 vs 沃特福德	主胜 1-1 [51.1/25.4/23.5]	主胜 1-1 [49.3/25.7/25.0]	一致	一致
西汉姆联 vs 狼队	主胜 2-1 [45.1/25.7/29.2]	主胜 1-1 [44.8/26.4/28.8]	一致	不一致（弱）
伯明翰 vs 南安普敦	客胜 1-2 [37.0/25.8/37.2]	主胜 1-1 [38.8/25.8/35.4]	★ 不一致	不一致
斯托克城 vs 诺维奇	客胜 1-2 [29.8/25.1/45.1]	客胜 1-1 [30.8/25.7/43.4]	一致	不一致

概率为 [主胜 / 平 / 客胜]，%。盲测三项与原分析的差全部 ≤2pp——两边读的是同一张盘，差异不在基线而在基线之后。

分歧 ①：胜负翻转 1 场——伯明翰 vs 南安普敦

- 市场怎么说：独赢 2.39 / 3.60 / 2.62，去水后主 38.8 / 客 35.4；让球盘「伯明翰 0 @0.89 vs 南安普敦 0 @0.99」同样把主队定为小热门。两口径自治。
- 原分析怎么做：把「南安普敦因 Spygate 被罚 4 分、盘口排名 [5] 对应真实战力更强」当作利客证据，做了主 -1.5pp / 客 +1.5pp 的修正——数值上合规（<45% 档 ≤1.5pp），结果把 3.4pp 的主队优势翻成 0.2pp 的客队优势，结论从主胜变客胜。
- 盲测为什么不修：罚分 8 月初已公开，页头 [5] 的排名本身就是按真实战力开的，**这是已定价信息，再拿来修正就是双重计算**——正是方法论"反向纪律"要拦的情形。所以保持市场值，报主胜 38.8%。
- 更重要的教训：在市场差距只有 3pp 的均势盘上，一次"合规"的 1.5pp 修正就足以翻转结论。这说明"≤1.5pp"对这类场次并不是安全阈值——**当主客差 < 3pp 时，任何修正都可能翻方向，正确做法是报「无倾向」**，而不是让阈值来背锅。

分歧 ②：首选比分——"方向内比分"与"网格众数"是两套口径

8 场里 6 场的网格众数是 **1-1**：λ 全落在 2.6–2.9、主客拆分接近，右偏分布下 1-1 恒是众数。原分析则系统地报"与胜负方向一致的比分"（主胜报 2-1、客胜报 1-2）。这本身是规则 4 允许的分歧（众数比分与边缘概率可以分居两侧），**但必须显式标注**。问题在于——其中 3 场即便退一步、只看"预测方向内的众数"，原报的比分也不是最高的：

场次	λ 拆分（主/客）	方向内众数	原报	差距
林肯城 vs 布莱克本	1.41 / 1.23	1-0 = 10.1%	2-1 = 8.7%	报了方向内第 2 位
普雷斯顿 vs 布里斯托城	1.28 / 1.39	0-1 = 9.6%	1-2 = 8.6%	报了方向内第 2 位
斯托克城 vs 诺维奇	1.25 / 1.53	0-1 = 9.5%	1-2 = 9.1%	报了方向内第 2 位

另外 3 场（谢菲尔德联 1-0≈2-1、西汉姆联 2-1、伯明翰若按其客胜口径则 1-2）原报的确是方向内众数，只是低于全局众数 1-1，记作"弱不一致"。这类错误肉眼完全看不出来，只有把网格算出来才抓得到——这就是今天把「网格众数排序」并入 `check_consistency.py` 的原因，也是要求今后每场把 λ 拆分写进台账的原因。

大小球 / 双方进球：无一场方向相反，只有"该不该报方向"的差异

- **大小球**：8 场去水大球概率 47.5%–55.8%，与原分析的数字差都在 2pp 内。原报"大 2.5 (54%)"一类方向，按新规则（45–55% 一律「无倾向」）只有西汉姆联 (55.8%) 够得上报「大」。
- **双方进球**：泊松 51%–58%。原 8/8 全报「是」；盲测只有西汉姆联 58%、伯明翰 57%、斯托克城 56% 报「是」，其余 5 场「无倾向」。没有一场盲测会反过来报「否」。

批次 5（今晚 2 场）：全部一致

场次	原预测	盲测	判定
皇家社会 vs 塞尔塔	主胜 1-1 [50.0/25.5/24.5] · 大小无倾向 51% · BTTS 无倾向 53%	主胜 1-1 [50.3/25.6/24.0] · 无倾向 51% · 无倾向 53%	四口径全同
图卢兹 vs 里尔	客胜 1-1 [27.5/27.0/45.5] · 小 2.5 微倾 47.7% · BTTS 无倾向 52%	客胜 1-1 [27.0/27.0/46.1] · 无倾向 48% · 无倾向 52%	方向全同（大小球标签差）

- 三项概率差都 ≤0.6pp，胜负、比分、两个边缘口径全部相同。
- **赛前检索到的信息全是已定价的**：皇家社会缺 Odriozola（长期）、Marín 腿筋存疑，塞尔塔缺 Borja Iglesias，Matarazzo 停赛不能站场边；图卢兹缺 Francis / Nianzou / Igamane，Fernández-Pardo 已转会纽卡斯尔；里尔上轮差几秒赢巴黎。这些都是开赛前多日的公开新闻，按反向纪律不构成修正理由，盲测维持零修正。
- 唯一的口径差是图卢兹场的"小 2.5 微倾"——47.7% 按新规则应标无倾向，方向没变。
- 皇家社会场的截图里没有双方进球盘，BTTS 只用泊松值。

由此落地的三条改动

1. **网格众数排序校验**已并入发布闸门脚本：首选比分若不是网格众数即报 X；TOP3 顺序与网格不一致报 △。前提是台账写入 lam_home / lam_away。
2. **主客差 < 3pp 的均势盘**：不再依赖"≤1.5pp"阈值，直接报「无倾向」——伯明翰场证明阈值内的修正也能翻结论。
3. **大小球 / BTTS 落在 45–55% 一律「无倾向」**；同批全部同向视为未做逐场区分。

来源：Sports Mole：皇家社会 vs 塞尔塔 预览 · Yahoo Sports：皇家社会 vs 塞尔塔 预览 · Sports Mole：图卢兹 vs 里尔 预览 · Sportsgambler：图卢兹 vs 里尔。盘口数据来自用户投递的 m617 系截图（NAS 场次 014–023）。

本页为方法论复核记录，非投注建议。