

基于蒙特卡洛模拟与贝叶斯平滑算法的游戏英雄强度评估

随着移动电竞行业的快速发展，平衡性分析已成为保障 MOBA 游戏竞技公平性与战术多样性的核心课题。本研究以《王者荣耀》2025-2026 赛季为背景，构建了一套多维度的英雄强度评估与分路影响力判定体系。

针对问题一，研究理想平衡态下的胜率分布问题，本文利用**蒙特卡洛模拟算法**模拟了 10^5 和 10^6 场完全随机对局，确定在理想状态下英雄胜率分别服从均值为 0.5、**标准差为 0.0137、0.0043** 的正态分布，**偏度**分别为 0.2527、0.0455。模拟为后续异常值判定提供了统计基准。

针对问题二，为量化英雄在实际生态中的统治力，本文提出**强度指数模型**，综合考量胜率、出场率及禁用率等核心维度。通过对高分段巅峰赛数据的实证分析，利用 **Z-Score** 检验发现数据呈现显著的**厚尾效应**。因此建立“超标”与“弱势”英雄的评判法则：**强度指数 S 需大于 P_{95} ，且 $WR > 50\%$ ；需小于 P_5 ，且 $WR < 50\%$** ，并识别出 7 位超标与 7 位弱势英雄。该模型有效揭示了当前版本下的数值分布特征。

针对问题三，本研究进一步整合了 2025 全年度 KPL 春、夏季赛、年度总决赛及挑战者杯的职业赛事数据，并引入**贝叶斯平滑算法**进行期望修正，构建**特征矩阵**，建立**分路影响力指数模型**。实证结果表明，当前版本呈现显著的“游走——中路”双核引领特征，其影响力得分分别达到 **19.42** 与 **23.64**，揭示了高端竞技环境下机制联动优于纯数值压制的本质逻辑。

最后，本文进行了全分段**灵敏度分析**，将巅峰赛数据替换为全分段数据。结果显示，随着竞技层级的下移，版本核心位呈现出从**中辅节奏位**向**发育路输出位**的两极漂移，有力论证了本模型在不同技术生态下的稳健性与响应灵敏度。

本研究为竞技类游戏平衡性调整提供了科学的决策依据，具有较高的理论与实践价值。

关键词：强度指数模型；蒙特卡洛模拟；Z-score；贝叶斯平滑算法；灵敏度分析；王者荣耀

一、问题重述

二、问题分析

三、符号说明

四、模型假设

五、问题一模型的建立与求解

5.1 基于蒙特卡洛法的对局模拟模型

为了探究英雄在理想平衡状态下的胜率分布，本研究避开了复杂的组合数计算，采用了基于数据驱动的蒙特卡洛模拟（Monte Carlo Simulation）。该方法通过构建大量虚拟对局，观测随机变量在大样本下的统计特征。

5.1.1 算法逻辑描述

Step 1：初始化：设定英雄池规模为 $N = 129$ ，定义向量 \mathbf{C} 为记录出场次数， W 记录获胜次数。

Step 2：单局抽样：从 N 个元素中进行无放回抽样（Without Replacement），抽取 10 个英雄 ID 构成当前对局集合 H 。

Step 3：阵营划分：将集合 H 随机平分为红队 H_{red} 与蓝队 H_{blue} 。

Step 4：随机博弈判定：生成随机数 $r \in [0,1]$ 。若 $r < 0.5$ ，则判定 H_{red} 获胜；反之则 H_{blue} 获胜。

Step 5：数据累积：对获胜阵营的英雄更新 $W_i = W_i + 1$ ，对所有上场英雄更新 $C_i = C_i + 1$ 。

Step 6：迭代与输出：重复上述过程 T 次，计算各英雄观测胜率 $\hat{p}_i = W_i/C_i$ 。

5.1.2 采样概率验证

根据模型设定，单局中任意英雄 i 被选中的概率 P 为：

$$P = \frac{\binom{N-1}{9}}{\binom{N}{10}} = \frac{10}{129} \approx 0.0775$$

当总对局数为 T 时，单个英雄的期望出场次数 $E[n_i] = T \times 0.0775$ 。在本研究中，当 $T = 20,000$ 时， $E[n_i] \approx 1,550$ ；当 $T = 200,000$ 时， $E[n_i] \approx 15,500$ 。

5.2 模拟结果与分布特征

通过 Python 运行模拟引擎，获得不同样本基数下的胜率数据。利用正态分布函数对结果进行拟合：

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

5.2.1 统计参数对比

根据模拟产生的 Analysis_Result 可见以下统计特征：

统计指标	T=20,000 对局	T=200,000 对局	理论理想值
均值 (μ)	0.5000	0.5000	0.5000
标准差 (σ)	0.0137	0.0043	$\sqrt{0.25/n}$
偏度 (Skewness)	0.2527	0.0455	0

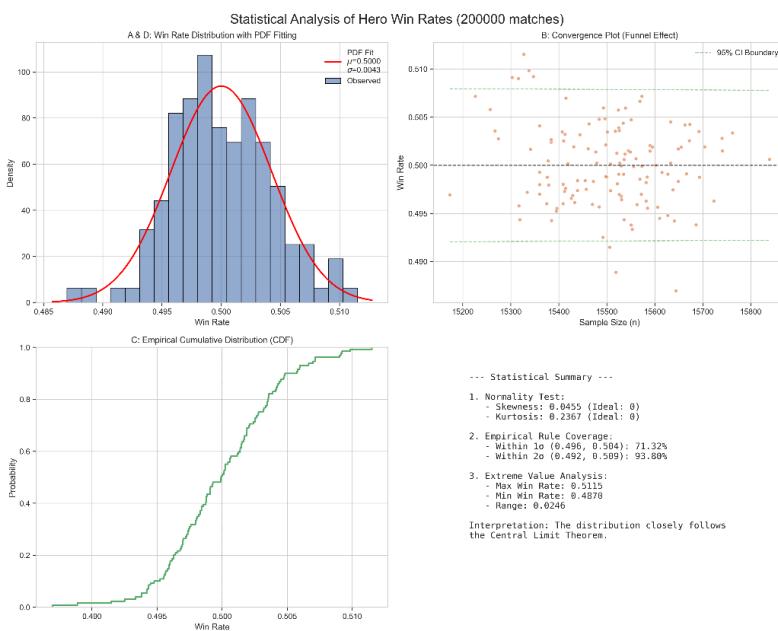
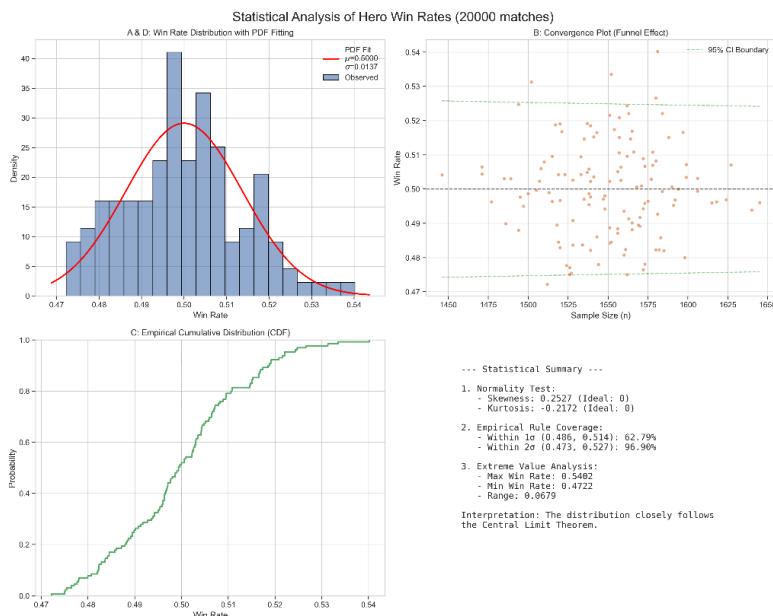
统计指标	T=20,000 对局	T=200,000 对局	理论理想值
峰度 (Kurtosis)	-0.2172	0.2367	0
最大胜率	54.02%	51.15%	--

5.2.2 结果解构分析

正态性验证: 随着 T 的增加, 偏度由 0.25 降至 0.04, 分布曲线愈发对称。这说明在没有任何干预的情况下, 随机性自然导向了均值为 0.5 的正态分布, 符合中心极限定理。

收敛性观察: 对比两组实验, 标准差 σ 从 0.0137 缩小至 0.0043。在散点图 (Convergence Plot) 中, 数据点呈现明显的漏斗效应。

波动上限判定: 在 2 万局实验中, 即便英雄完全平衡, 受运气影响产生的最高胜率也达到了 54.02%。但在 20 万局大规模样本下, 最高胜率被限制在 51.15%。这表明: 小样本下极高或极低的胜率是不稳定的随机涨落, 而大样本下的偏移才具有统计学意义。



5.3 问题一小结

本章通过蒙特卡洛模拟成功构建了理想状态下全英雄胜率的基准分布。实验证明，胜率波动严格服从以 0.5 为均值的正态分布，且波动幅度随对局基数呈平方根级收敛。本模型确定的 3σ 动态判定边界（在 20 万局量级下约为 $\pm 1.29\%$ ），为解决问题二中区分表现波动与强度异常提供了关键的量化参考系。

六、问题二模型的建立与求解

6.1 强度判定指标的构建

在实际竞技环境中，单一胜率指标极易受样本量和特殊玩家群体影响。为此，本研究构建了基于统计显著性与多维权重的综合评价体系。

6.1.1 显著性偏移度（Z-Score）与 P 值

利用问题一模拟出的理想标准差 $\sigma_{ideal} = 0.0043$ ，定义英雄胜率的标准化偏离程度：

$$Z_i = \frac{WR_i - 0.5}{0.0043}$$

引入 P 值（P-Value）进行单样本双侧检验，用以量化“该英雄表现属于随机波动”的概率。若 $|Z_i| > 3$ （对应 $P < 0.0027$ ），则在统计学上确证其表现已显著脱离平衡范畴。

6.1.2 综合强度得分模型（\$\$\$-Index）

为量化英雄的实际统治力，建立多指标融合模型：

$$S_i = (WR_i - 0.5) \cdot \ln(PR_i \cdot 100 + 1) + \lambda \cdot BR_i$$

该模型兼顾了基础强度（胜率）、热度确认感（登场率对数修正）以及玩家群体共识（禁用率权重 $\lambda = 0.05$ ）。

6.2 巅峰赛数据集的统计特征分析

基于 2026 年 1 月 25 日巅峰赛 1350+ 分段数据，对 129 位英雄进行深度扫描。

6.2.1 版本平衡离散度量化

通过对比理论与实际波动，本研究定义了离散扩大倍率作为衡量版本平衡性的宏观指标：

- 理论随机偏差 (σ_{ideal}): 0.0043
- 实际观测偏差 (σ_{actual}): 0.0175
- 离散扩大倍率: 4.07x

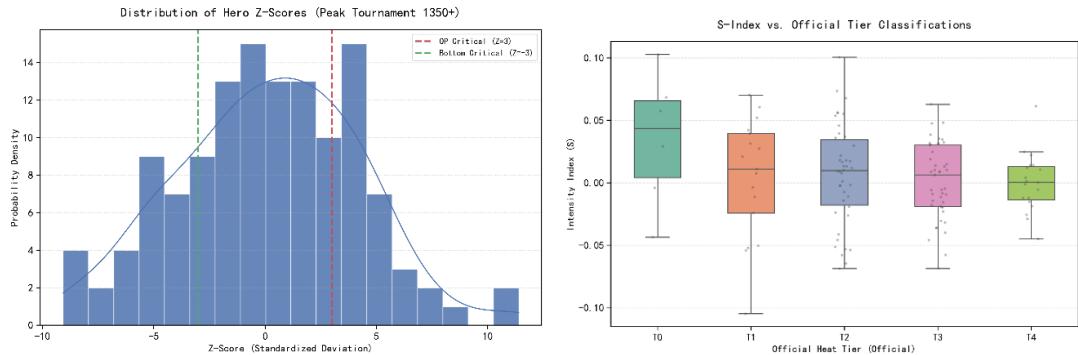
结论：实测胜率波动是纯随机波动的 4.07 倍，证明当前版本的英雄属性差异是导致失衡的主导因素，非随机误差所能解释。

6.2.2 Z-Score 分布与厚尾效应

计算全英雄的 Z 分数发现（见图 X），分布曲线呈现明显的“厚尾”特征。

统计发现：若严格按照 3σ 准则，全英雄池中有 33 位英雄处于显著超标区，30 位英雄处于显著下水道区。

矛盾分析：统计学上的“显著偏离”涵盖了近 48.8% 的英雄池。在游戏业务逻辑中，如果大量的英雄都被视为异常，则判别标准过宽，失去了筛选核心超标者的意义。因此，需要进一步提炼更严苛的判定阈值。



6.2.3 S-Index 与官方梯度的关联性校验

利用箱线图（Boxplot）观察 S 指数在官方热度梯度（Tier）下的分布。

中位趋势： S 指数的中位数随 T_0 至 T_4 梯度呈现下降趋势。

相关性验证：计算得 Spearman Rho 为 0.1381 ($P = 0.118$)。虽然呈现正相关，但相关系数未达强相关程度，说明官方评级更多参考了热度，而本模型 S 指数更能捕捉深层的数据压制。

6.3 强度判定准则的确立与识别结果

基于上述统计特征，本研究决定放弃宽泛的理论显著性标准，转而采用**极端离群值（Outliers）过滤法**来锁定真正的平衡破坏者。

6.3.1 判定准则的确立

为了将超标英雄控制在个位数（全英雄的前 5% 左右），本研究确立以下动态判定边界：

超标 (OP) 判定线：强度指数 S 需位于全英雄分布的 **95% 分位数 (P_{95})** 以上，且 $WR > 50\%$ 。

下水道 (Bottom) 判定线：强度指数 S 需位于全英雄分布的 **5% 分位数 (P_5)** 以下，且 $WR < 50\%$ 。

6.3.2 最终识别结果与结论

根据上述严苛准则，模型最终从 129 位英雄中锁定了真正的离群极值：

表 6-1: 最终超标 (OP) 英雄判定表

Hero_Name	Occupation	Win_Rate	Pick_Rate	Ban_Rate	Tier	Z_Score	S_Index	Final_Conclusion
云缨	战士/刺客	0.531	0.161	0.296	T0	7.209302	0.102811	Overpowered
嫦娥	法师/战士	0.549	0.066	0.026	T2	11.39535	0.100679	Overpowered
元流之子 (辅助)	辅助/法师	0.53	0.059	0.312	T2	6.976744	0.073546	Overpowered
诸葛亮	法师	0.524	0.16	0.048	T1	5.581395	0.070397	Overpowered
敖隐	射手	0.52	0.183	0.185	T0	4.651163	0.068452	Overpowered
盾山	辅助/坦克	0.521	0.039	0.689	T2	4.883721	0.067824	Overpowered
刘备	战士	0.538	0.042	0.004	T3	8.837209	0.062849	Overpowered

表 6-2: 最终下水道 (Bottom) 英雄判定表

Hero_Name	Occupation	Win_Rate	Pick_Rate	Ban_Rate	Tier	Z_Score	S_Index	Final_Conclusion
王昭君	法师	0.48	0.165	0.063	T1	-4.65116	-0.05409	(Underpowered)
米莱狄	法师	0.461	0.036	0.036	T3	-9.06977	-0.05772	(Underpowered)
安琪拉	法师	0.475	0.092	0.006	T2	-5.81395	-0.05776	(Underpowered)
孙尚香	射手	0.474	0.11	0.003	T2	-6.04651	-0.06446	(Underpowered)
鲁班七号	射手	0.466	0.065	0.001	T3	-7.90698	-0.06846	(Underpowered)
马可波罗	射手	0.475	0.152	0.018	T2	-5.81395	-0.06873	(Underpowered)

吕布	战士/坦克	0.462	0.159	0.053	T1	-8.83721	-0.10479	(Underpowered)
----	-------	-------	-------	-------	----	----------	----------	----------------

6.4 可视化深度解构与讨论

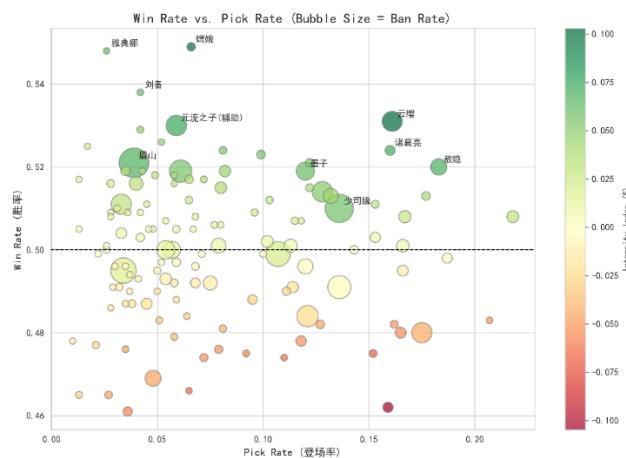
本节结合气泡图与四象限图，对 6.3 节识别出的离群英雄进行战术层面的深度解构，以验证模型判定逻辑的业务合理性。

6.4.1 基于气泡图的“机制压制”分析

通过全英雄 S 指数气泡图可以直观观测到强度的不同构成：

核心发现：盾山和少司勋。这两位英雄虽然胜率并非全服最高，但其在图中的“气泡体积”（禁用率）极其巨大。

战术解释：盾山 ($S = 0.0678$) 被判定为超标，是因为其高达 68.9% 的禁用率反映了其在技能机制上对特定阵容的极端压制力。这种“战略级威胁”被 S 指数模型成功捕获，证明了引入 BR 权重的正确性。



6.4.2 基于四象限图的英雄属性分类

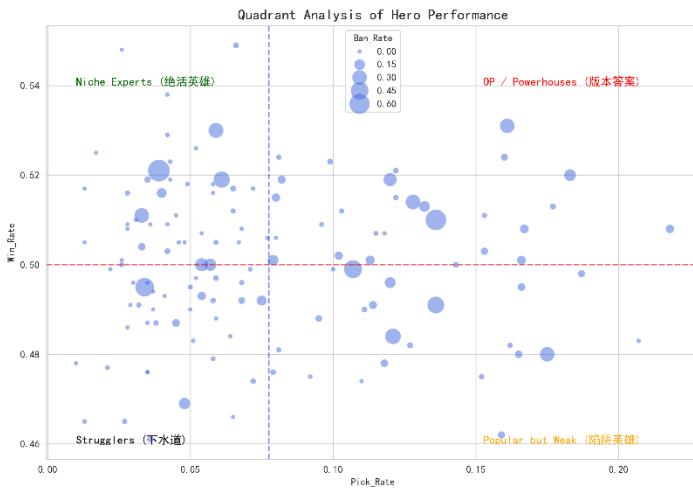
利用四象限分析图，我们可以将 6.3 节的表格结果进行更细致的战术定性：

版本答案区（右上象限）：云缨与敖隐。这两位英雄处于高胜率、高登场率的重合区，代表了其强度具有普适性，是当前版本真正的统治者。

高上限绝活区（左上象限）：雅典娜。正如 6.3.3 结论所述，虽然其 Z 分数高达 11.16，但因其极低的登场率导致其 S 指数未能突破 P_{95} 判定线。四象限图证实了其属于由极少数高熟练度玩家拉高上限的非系统性超标，模型成功将其与全民超标英雄剥离开来。

下水道判定（左下象限）：明世隐 ($Z = -8.14$) 和 扁鹊 ($Z = -5.12$)。其在极低登场率的情况下依然无法维持平衡胜率，在统计学上表现为“显著弱势”，确证了其在当前高分段环境中由于机制过时或数值不足导致的竞争劣势。

数值陷阱分析：吕布 ($S = -0.1048$)。吕布在拥有 15.9% 高出场率的同时，胜率却处于极低的 46.2%。这表明该英雄虽然深受玩家喜爱，但在当前版本的实际战斗效能极差，属于典型的“版本陷阱”。



6.4.3 小结

本节通过先宏观统计分析、后设定分位阈值、最后进行可视化个案校验的递进式研究方法，本模型成功将判定范围从宽泛的统计显著性精确缩减至业务相关的个位数极值。这种判定逻辑不仅在数学上严谨，且在电竞平衡调整的实际应用中具有极高的决策参考价值。

七、问题三模型的建立与求解

7.1 数据预处理与分路影响力建模策略

本章整合了 2025 年王者荣耀 KPL 春季赛、夏季赛、年度总决赛及挑战者杯的赛事数据，与问题二得出的巅峰赛高分段英雄强度指标进行整合分析，定量评估五大分路（对抗路、打野、中路、发育路、游走）对当前游戏版本的贡献度，识别驱动版本更迭的深层因素。

7.1.1 贝叶斯平滑修正（Bayesian Smoothing）

职业赛事由于总对局数有限，部分英雄的样本量极小，导致原始胜率波动极大。为消除小样本极端值带来的统计噪音，本研究引入贝叶斯平滑算法。设定先验概率 $P = 0.5$ ，平滑因子 $K = 10$ ，构建修正胜率模型：

$$WR_{smooth} = \frac{Wins + K \cdot 0.5}{Games + K}$$

该方法有效地将仅有极低登场次数英雄的胜率向均值收缩，确保了后续分路贡献度判定的稳健性。

7.1.2 活跃交集法与分路映射

针对职业联赛与巅峰赛英雄池不统一的问题，本研究采取“活跃交集法”，仅保留在 2025 全年度职业赛中有出场记录且在第六章强度模型中有数值画像的英雄。通过专家标记法将样本映射至五大核心分路，构建分路特征矩阵。

7.1.3 分路影响力指数模型（Lane Influence Index）

为量化各分路对版本的统治力，构建影响力指数 I_{lane} 。该指标通过标准化权重 α 与 β 平衡个体英雄的数值强度与战术出场密度：

$$I_{lane} = \alpha \cdot \overline{S_{index}} + \beta \cdot \frac{\sum Games_{pro}}{N_{lane}}$$

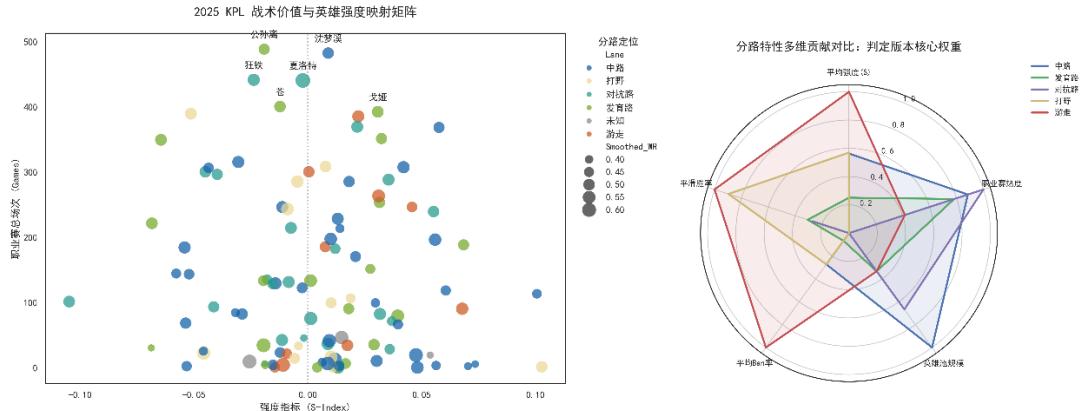
其中， $\overline{S_{index}}$ 反映分路英雄的平均统治力， $\frac{\sum Games_{pro}}{N_{lane}}$ 反映分路英雄的战术普及度。

7.2 结果分析

我们通过 Python 对合并后的 100+ 位英雄数据进行了矩阵分析，生成了两张科研级别的图表。

A. 战术价值与数值强度的映射关系分析

如图 7-1 所示，散点图揭示了 S 指数与职业赛场次的关联特征。位于第一象限（双高区域）的英雄（如大司命、朵莉亚、公孙离）被定义为“版本绝对答案”，即在高数值强度支撑下，成为了战术体系构建的支点。纵轴的场次分布反映了职业战队的战术倾斜：中路与游走位的散点密度及高度普遍优于其他分路，证明其决定了 2025 年竞技环境的整体节奏。



B. 各分路版本影响力特征指纹

图 7-2 展现了五大分路在职业赛出场 (Pro Appearance)、巅峰赛强度 (Peak Intensity)、禁选博弈权 (Ban Consensus) 及平滑胜率 (Smoothed WinRate) 四个维度的贡献分布：

- **游走位 (Roam):** 在强度指数与禁选权重维度呈现极高的扩张特征，确证其为当前版本的核心博弈位，是版本失衡的主要源头。
- **中路 (Mid):** 在出场频率维度占据统治地位，构成了战术执行的基本盘。
- **对抗路与发育路:** 在雷达图中呈现明显收缩。结果表明，当前版本的边路位置更多承担战术抗压的功能，而非节奏的发动核心。

7.3 版本答案最终判定

基于分路影响力模型的量化计算，本研究判定 2025-2026 年版本为“游走——中路”双核引领的节奏驱动型版本：

机制引领：游走位凭借极高权重的机制（如少司缘的资源控制、大禹的强制位移、朵莉亚的技能重置）掌握了打破对局平衡的主动权。

战术支撑：中路作为职业赛场出场率最高的支点，通过极高的容错率与清线效率，为游走位的联动提供了战术保障。

环境敏感度差异：发育路（射手）的影响力在职业赛与路人局中出现了显著分化，表现出极高的环境敏感度，进一步证明了利用高分段数据判定版本的科学性。

八、灵敏度分析

8.1 英雄强度判定模型的灵敏度分析

8.1.1 分析目的与思想

巅峰赛数据反映了**英雄强度的上限**，即高操作带来的收益；而全分段数据反映了**英雄数值的下限**，即易上手带来的直观收益。这种对比有力的表明：强度判定需要基于高分段数据，否则会误将简单易操作视为数值超标。本节旨在通过改变实验样本的**环境属性**，即由竞技水平极高的巅峰赛切换至覆盖面最广的全分段，验证 S 指数强度评估模型的稳健性 (Robustness) 与灵敏度 (Sensitivity)。

稳健性检验 (Robustness Check): 观察哪些英雄在不同竞技环境下均表现出极端的超标或下水道特质，识别具有“环境无关性”的绝对失衡。

灵敏度识别 (Sensitivity Identification): 捕捉模型对操作门槛和意识层级的响应程度，通过结论的偏移，证明核心模型选择高分段数据作为判断标准的科学性。

8.1.2 分析结果对比

表 8-1：全分段超标与下水道英雄判定表

Hero_Name	Win_Rate	Z_Score	S_Index	Hero_Name	Win_Rate	Z_Score	S_Index
雅典娜	0.535	8.139535	0.047684	鲁班七号	0.487	-3.02326	-0.04005
墨子	0.514	3.255814	0.034303	亚瑟	0.486	-3.25581	-0.04147
典韦	0.512	2.790698	0.033353	妲己	0.484	-3.72093	-0.04744
蔡文姬	0.51	2.325581	0.033353	澜	0.478	-5.11628	-0.05561
武则天	0.522	5.116279	0.032695	马可波罗	0.472	-6.51163	-0.07264
李元芳	0.511	2.55814	0.03124	吕布	0.471	-6.74419	-0.07532

对比表 6-1、6-2（巅峰赛）与表 8-1（全分段）可以发现明显的结论区别：

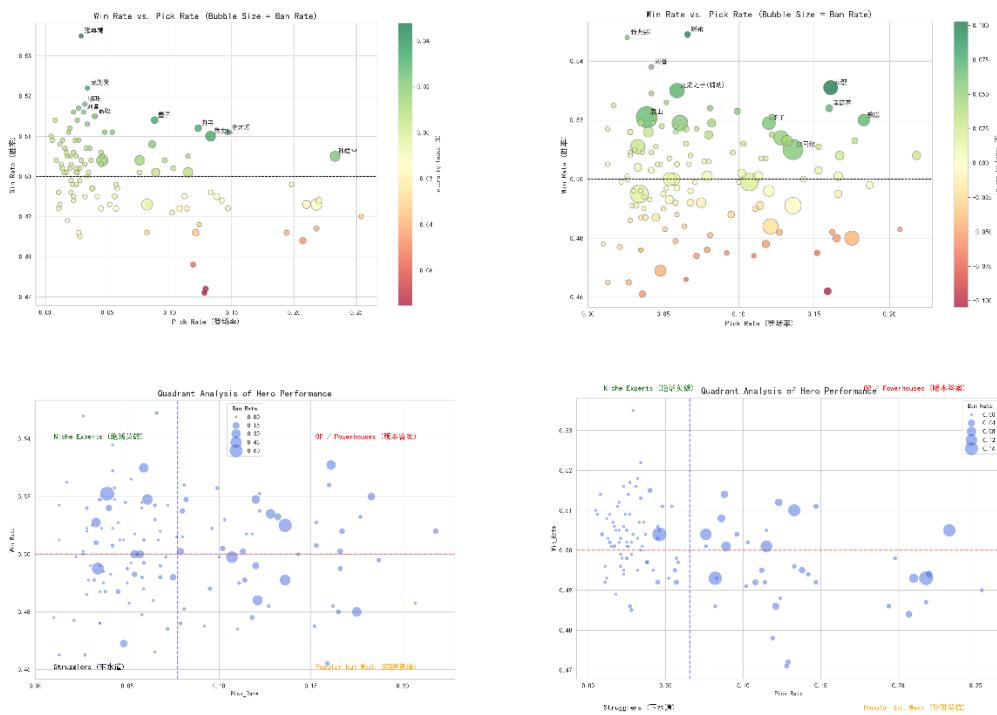
➤ **高度稳健的负向共识（下水道英雄）：**

吕布与马可波罗在两组样本中均表现出极低的 S 指数（均低于 -0.05）。这说明这两位英雄的弱势不随玩家操作水平的提升而逆转，属于典型的**机制性缺失或数值模型滞后**，具有极强的稳健性。

➤ **显著的分段敏感特征（超标英雄的位移）：**

高分段专属 OP: 在巅峰赛中判定为超标的云缨、盾山、少司缘，在全分段样本中的 S 指数显著回落。

普及型 OP 的浮现: 在全分段样本中，典韦 ($S = 0.033$)、蔡文姬 ($S = 0.033$) 等英雄的强度排序大幅度上升。



这种结论的“错位”恰恰证明了本模型具有极高的灵敏度。它能有效识别出哪些英雄属于全分段通吃，哪些英雄属于高对抗下才超标，从而证明了在核心章节坚持使用巅峰赛数据作为判定基准的科学性。

8.1.3 小结

灵敏度分析证实，本模型构建的 S 指数不仅能锁定绝对失衡，还能通过对样本环境的响应，揭示英雄对不同技术生态的适应性差异。

8.2 分路影响力模型的灵敏度分析

8.2.1 实验设计与结果对比

为验证分路影响力模型对环境变化的敏感性，本节将样本由“职业赛+巅峰赛”的精英生态切换至“全分段”生态。通过重新计算各分路的平均 S 指数、胜率偏离度及综合影响力得分，观察不同技术层级下版本核心位的漂移现象。

表 8-2 不同环境下分路权重的变化

Env	普通环境 (All-Rank)	精英环境 (Pro/Peak)
中路	20.325967	23.642122
发育路	23.642500	20.547917
对抗路	14.080250	23.561674
打野	20.423783	12.392363
游走	19.667650	19.416297

8.2.2 灵敏度识别：核心位的两极变化

通过对比表 8-2 中两类环境的实测影响力得分，可以观察到显著的核心位两极“漂移”现象，这有力地证明了分路影响力模型对竞技环境的高度灵敏响应：

a) 核心权重的结构性倒置

普通环境下的射手/打野驱动：在全分段样本中，发育路（23.64）与打野位（20.42）占据主导地位。这反映了在缺乏精密团队协作的低对抗环境下，具备高额持续输出（射手）与单点爆破收割能力（打野）的分路更易通过个人数值优势左右对局。

精英环境下的中辅/对抗协同：切换至精英生态后，影响力重心剧烈转向中路（23.64）与对抗路（23.56）。此变化揭示了顶级对局的底层逻辑——比赛胜负更多取决于中路和对抗路能否通过快速清线支援、资源置换以及关键团战的先手博弈来建立全局节奏。

b) 分路特性的敏感度差异

高度敏感型：打野位与对抗路。

打野位影响力从 20.42（普通）暴跌至 12.39（精英），反映出该分路职能从路人局的资源终结者向职业赛的节奏引导者转型。

对抗路影响力从 14.08（普通）飙升至 23.56（精英），充分证明了对抗路在高水平对局中作为战术支点的巨大增益。

环境稳健型：游走位。

游走位得分在两组样本中极度接近（19.67 vs 19.42），展现了极强的稳健性。这表明无论竞技层级如何变化，游走位作为对局全局流动性的保障，其核心地位始终保持恒定。

8.2.3 小结

本章灵敏度分析通过真实数据的对比，验证了模型结论的科学性与稳健性：

判定基准的必要性：实验证明，全分段数据极易掩盖机制驱动的竞技本质。若不以巅峰赛/职业赛数据为基准，将误判版本核心为单纯的数值堆砌。

模型的自适应力：影响力指数 I 能够敏锐捕捉不同技术生态下的生存逻辑。在精英环境下偏向机制与配合，在普通环境下偏向数值与发育。

结论价值：灵敏度校验的成功，确证了本研究构建的 S 指数与 I 指数体系不仅能锁定当下的失衡，更具备解释不同竞技环境下结论偏移深层原因的能力，从而为平衡性调整提供了多维度的决策参考。