

你好！我是MCM_C。

非常棒的提问。你不仅是在问“这两张图是什么”，而是在问**“如何围绕这两张图，构建这道题（Problem C Q2）的完整解题框架”**。这正是获得美赛Outstanding Winner（O奖）的关键思维——以图表为锚点，串联起严密的逻辑链条。

这道题的核心任务是机制对比（**Mechanism Comparison**）。为了回答“哪种机制更偏向观众”，你需要构建一个包含反演（**Inversion**）、**交叉验证（Cross-Evaluation）和灵敏度分析（Sensitivity Analysis）**的完整模型体系。

以下我将基于你提供的图2_1_1和图2_1_2，为你梳理这道题的完整建模方法论（**Complete Methodology**）。你可以直接参考这个结构来撰写你论文的对应章节。

论文章节规划：基于反事实推演的评分机制对比模型

1. 模型定义与机制数学化 (Formalization)

首先，必须用数学语言清晰定义两种“游戏规则”。这是对比的基础。

- 符号定义：
 - J_i : 选手 i 的评委原始分。
 - F_i : 选手 i 的观众原始票数（这是未知量，需反演）。
 - N : 参赛选手总数。
- 机制 A: 排名积分法 (The Rank Method)
 - 规则: 将评委分排名 $r_{J,i}$ 与观众票排名 $r_{F,i}$ 直接相加。
 - 公式: $S_{rank,i} = r_{J,i} + r_{F,i}$
 - 淘汰判据: S_{rank} 最小者（或最大者，视具体规则定义）淘汰。
 - 数学特性: 离散化，丢弃了“分差”信息（即：赢1分和赢10分效果一样）。
- 机制 B: 百分比积分法 (The Percentage Method)
 - 规则: 将评委分占比与观众票占比相加。
 - 公式: $S_{pct,i} = w \cdot \frac{J_i}{\sum J} + (1-w) \cdot \frac{F_i}{\sum F}$
 - 数学特性: 连续化，保留了“分差”信息，且引入了权重参数 w （通常默认为50/50）。

2. 步骤一：数据重构 (The Inverse Problem)

对应你的图 2_1_1 (fileName: 2_1_1.png)

- 逻辑痛点: 题目问“对比两种方法在各赛季的结果”。但我们只有评委分和最终淘汰结果，没有观众具体的投票数。没有 F_i ，就无法计算机制B的结果。
- 解决方法: 蒙特卡洛反演 (Monte Carlo Inversion) 或 极大似然估计 (MLE)
 - 核心思想: 既然我们知道谁被淘汰了，我们可以反推“当时的观众投票大概长什么样，才会导致这个结果”。
 - 操作流程:
 - 设定观众投票服从某种分布（如Pareto分布或Zipf分布，模拟头部效应）。

- 2. 生成成千上万组随机的观众投票数据。
- 3. 筛选出那些能完美复现历史淘汰结果的数据组。
- 4. 对筛选出的数据取平均, 得到***“最可能的历史观众投票分布”**。
- 图8在文中的地位: 它是这一步的产出验证。展示了你通过模型还原出的“历史真相”(例如 Jerry Rice 极高的观众支持率), 证明你的数据基础是扎实的。

3. 步骤二: 交叉验证与反事实推演 (Counterfactual Simulation)

这是回答题目 “Apply both approaches to each season” 的核心步骤。

- 核心逻辑:
 - 对于 Season 1-2 (原本用排名法): 保持评委分和反推的观众票不变, 强行套用“百分比法”, 看谁会晋级/淘汰。
 - 对于 Season 3+ (原本用百分比法): 同理, 强行套用“排名法”, 看结果有何不同。
- 产出指标:
 - 排名一致性 (Rank Concordance): 两种方法得出的排名列表, Kendall's Tau 相关系数是多少?
 - 翻盘事件 (Overturn Event): 统计有多少次“评委打分低但观众投票高”的选手, 在一种机制下活了, 在另一种机制下死了。

4. 步骤三: 偏差量化与灵敏度分析 (Bias Quantification)

对应你的图 2_1_2 (fileName: 2_1_2.png)

- 逻辑痛点: 仅仅算出“结果不同”还不够, 题目问的是“是否更偏向观众 (Favor Fan Votes)”。
- 解决方法: 参数扫描 (Parameter Sweep)
 - 变量: 调节观众投票的权重 w_f (或者理解为调节“观众狂热度”的影响因子)。
 - 观测对象: 翻盘率 (Overturn Rate) —— 定义为“评委排名靠后, 但依靠观众投票成功晋级”的概率。
- 图9在文中的地位: 它是最终结论的铁证。
 - 解读逻辑:
 - 我们在图2_1_2中看到, 红色曲线(百分比法)始终位于蓝色曲线(排名法)之上, 且斜率更陡。
 - 这意味着: 在相同的观众支持度下, 百分比法能让选手以更高的概率“逆袭”评委的打分。
 - 结论: 数学上证明, 百分比法是一个***“民意放大器” (Amplifier), 而排名法是一个“平滑器” (Equalizer)**。因此, 百分比法更偏向观众。

总结: 如何把这个流程写进论文?

建议你的论文结构如下:

1. 4.1 The Mechanism Logic: 列出公式, 对比 Rank vs Percentage 的数学性质 (离散 vs 连

续)。

2. **4.2 Reconstructing History (Inverse Modeling):** 描述你怎么用图2_1_1的方法反推了 Jerry Rice等人的票数。
3. **4.3 The "What-If" Experiment:** 描述交叉验证的过程。
4. **4.4 Sensitivity & Bias Analysis:** 放上图2_1_2。详细论述“S型曲线”意味着百分比法对观众投票更敏感。
5. **4.5 Conclusion:** 回答问题——“Yes, the Percentage Method inherently favors fan votes due to its sensitivity to extreme values (distribution skewness), whereas the Rank Method dampens this effect.”

这一套方法论结合了数据挖掘(反演)、仿真模拟和系统分析,是非常标准的数学建模竞赛高分套路。