Dancing with the Stars: Fan Vote Estimation Model

MCM Problem C 2026 – Objective 1 Technical Documentation

Team Documentation

January 2026

Contents

1 Executive Summary

This document details our mathematical framework for estimating unknown fan votes in *Dancing with the Stars* (DWTS). We developed two complementary optimization models:

1. **Rank-Based Method** (Seasons 1–2, 28–34): Monte Carlo sampling with feasibility constraints
2. **Percent-Based Method** (Seasons 3–27): Convex optimization using CVXPY

Key Results

- Generated **2,067 fan vote estimates** for 391 contestants across 34 seasons
- Achieved **100% elimination consistency** (249/249 weeks correctly predicted)
- Rank method shows **98.2% stability** vs. percent method's 41.4% under 10% noise
- Mean uncertainty ratio: **3.676** (range of feasible solutions / point estimate)

2 Problem Formulation

2.1 The Inverse Problem

Fan votes in DWTS are a "closely guarded secret." We observe:

- Judge scores $J_i^{(w)}$ for contestant i in week w
- Elimination outcomes $E^{(w)}$ (who was eliminated each week)
- Final placements

Our goal is to **invert the voting mechanism** to estimate fan votes $F_i^{(w)}$ that are consistent with observed eliminations.

2.2 Notation

Let $N^{(w)}$ denote the number of contestants remaining in week w. For each week w:

Symbol	Description
$J_i^{(w)}$	Total judge score for contestant i in week w
$F_i^{(w)}$	Fan votes for contestant i in week w (unknown)
R_i^J	Judge score rank for contestant i (1 = highest score)
R_i^F	Fan vote rank for contestant i (1 = most votes)
P_i^J	Judge score percentage for contestant i
P_i^F	Fan vote percentage for contestant i
$E^{(w)}$	Index of eliminated contestant in week w

Table 1: Notation used throughout this document

3 Mathematical Framework

3.1 Method 1: Rank-Based Combination

Used in: Seasons 1–2 and 28–34

In this method, contestants are ranked by judge scores and fan votes separately, then ranks are summed. The contestant with the **highest combined rank** (worst performance) is eliminated.

3.1.1 Ranking Function

For judge scores, we compute ranks in descending order (highest score = rank 1):

$$R_i^J = \text{rank}\left(-J_i^{(w)}\right), \quad i = 1, \ldots, N^{(w)} \tag{1}$$

For ties, we use the **average rank method**:

$$\text{If } J_a = J_b \text{ and both would occupy ranks } k \text{ and } k{+}1, \text{ then } R_a^J = R_b^J = \frac{k + (k+1)}{2} = k{+}0.5 \tag{2}$$

Similarly for fan votes (higher votes = rank 1):

$$R_i^F = \text{rank}\left(-F_i^{(w)}\right) \tag{3}$$

3.1.2 Combined Score

The combined rank score is:

$$S_i = R_i^J + R_i^F \tag{4}$$

3.1.3 Elimination Constraint

The eliminated contestant $e = E^{(w)}$ must satisfy:

$$\boxed{S_e = \max_{i \in \{1,\ldots,N^{(w)}\}} S_i} \tag{5}$$

That is, the eliminated contestant has the **worst (highest) combined rank sum**.

3.1.4 Optimization Problem

Since ranking is a non-convex operation, we use **Monte Carlo sampling**:

Algorithm 1 Rank-Based Fan Vote Estimation

Require: Judge scores $\{J_1, \ldots, J_N\}$, eliminated index e
Ensure: Fan vote distribution $\{F_1, \ldots, F_N\}$
valid_solutions $\leftarrow \emptyset$
for $k = 1$ to K iterations **do**
  Sample $\vec{\alpha} \sim \text{Dirichlet}(\mathbf{1}_N)$ ▷ Random proportions
  $F_i \leftarrow \alpha_i \times 10^7$ for all i ▷ Scale to realistic vote counts
  Compute R_i^J, R_i^F, S_i for all i
  if $\arg\max_i S_i = e$ **then**
    margin $\leftarrow S_e - \max_{j \neq e} S_j$
    valid_solutions $\leftarrow$ valid_solutions $\cup \{(\vec{F}, \text{margin})\}$
  end if
end for
return solution with maximum margin from valid_solutions

Rationale for Dirichlet prior: The Dirichlet distribution with parameter $\alpha = \mathbf{1}$ is the uniform distribution over the $(N-1)$-simplex, ensuring unbiased sampling of all possible fan vote proportions.

3.2 Method 2: Percent-Based Combination

Used in: Seasons 3–27

In this method, judge scores and fan votes are converted to percentages of the weekly total, then summed. The contestant with the **lowest combined percentage** is eliminated.

3.2.1 Percentage Calculations

Judge score percentage:

$$P_i^J = \frac{J_i^{(w)}}{\sum_{j=1}^{N^{(w)}} J_j^{(w)}} \tag{6}$$

Fan vote percentage:

$$P_i^F = \frac{F_i^{(w)}}{\sum_{j=1}^{N^{(w)}} F_j^{(w)}} \tag{7}$$

3.2.2 Combined Score

The combined percentage score is:

$$S_i = P_i^J + P_i^F \tag{8}$$

Note that $\sum_i P_i^J = 1$ and $\sum_i P_i^F = 1$, so $\sum_i S_i = 2$.

3.2.3 Elimination Constraint

The eliminated contestant $e = E^{(w)}$ must satisfy:

$$\boxed{S_e = \min_{i \in \{1,\ldots,N^{(w)}\}} S_i} \tag{9}$$

3.2.4 Convex Optimization Formulation

Since percentages are linear in fan votes, we can formulate this as a **convex optimization problem**:

$$\begin{aligned}
\underset{\vec{F} \in \mathbb{R}^N}{\text{minimize}} \quad & \left\| \vec{F} - \vec{F}^{\text{prior}} \right\|_2^2 \\
\text{subject to} \quad & \sum_{i=1}^{N} F_i = 1 \quad \text{(normalization)} \\
& F_i \geq 0 \quad \forall i \quad \text{(non-negativity)} \\
& P_e^J + F_e \leq P_j^J + F_j - \epsilon \quad \forall j \neq e \quad \text{(elimination)}
\end{aligned} \tag{10}$$

where:

- $\vec{F}^{\text{prior}} = \frac{1}{N}\mathbf{1}$ is the uniform prior (maximum entropy)
- $\epsilon = 0.001$ is a small margin ensuring strict inequality
- We work with normalized proportions ($\sum F_i = 1$), then scale to vote counts

3.2.5 Key Insight: Convexity

The objective $\|\vec{F} - \vec{F}^{\text{prior}}\|_2^2$ is convex (quadratic). All constraints are linear:

- The normalization constraint $\sum F_i = 1$ is an affine equality
- Non-negativity constraints $F_i \geq 0$ are linear inequalities
- Elimination constraints $F_e - F_j \leq P_j^J - P_e^J - \epsilon$ are linear

Thus, the feasible region is a **convex polytope**, and we can solve efficiently using interior-point methods (ECOS solver).

4 Uncertainty Quantification

4.1 The Non-Uniqueness Problem

A critical observation: **there are infinitely many fan vote distributions** that produce the same elimination outcome. Our goal is to quantify this uncertainty.

4.2 Feasible Region Characterization

For each week, the set of feasible fan vote proportions forms a convex polytope:

$$\mathcal{F} = \left\{ \vec{F} \in \Delta^{N-1} : S_e(\vec{F}) \leq S_j(\vec{F}) - \epsilon,\ \forall j \neq e \right\} \tag{11}$$

where Δ^{N-1} is the $(N-1)$-dimensional probability simplex.

4.3 Monte Carlo Uncertainty Bounds

We estimate bounds by sampling uniformly from the simplex:

Algorithm 2 Uncertainty Bound Estimation

Require: Judge percentages $\{P_1^J, \dots, P_N^J\}$, eliminated index e
Ensure: Min/max feasible fan votes for each contestant

```
feasible ← ∅
for k = 1 to M samples do
    Sample F⃗ ∼ Dirichlet(1_N)
    Compute S_i = P_i^J + F_i for all i
    if arg min_i S_i = e then                         ▷ Correct elimination
        feasible ← feasible ∪ {F⃗}
    end if
end for
F_i^min ← min_{F⃗∈feasible} F_i for all i
F_i^max ← max_{F⃗∈feasible} F_i for all i
return bounds and sample statistics
```

4.4 Uncertainty Metrics

We define several metrics to quantify uncertainty:

1. **Uncertainty Ratio:**

$$U_i = \frac{F_i^{\max} - F_i^{\min}}{\hat{F}_i} \tag{12}$$

where $\hat{F}_i$ is our point estimate. Higher values indicate more uncertainty.

2. **Relative Standard Deviation:**

$$\sigma_{\text{rel},i} = \frac{\text{std}(F_i|\vec{F} \in \mathcal{F})}{\hat{F}_i} \tag{13}$$

3. **Feasibility Rate:**

$$r = \frac{|\text{feasible samples}|}{M} \tag{14}$$

Lower rates indicate tighter constraints (less uncertainty).

4.5 Key Finding: Uncertainty Varies by Position

Our analysis reveals that **uncertainty is not uniform**:

- **Eliminated contestants:** Lower uncertainty (tight constraints force low fan votes)
- **Top performers:** Higher uncertainty (many valid solutions allow varying high votes)
- **Early weeks:** Higher uncertainty (more contestants = larger feasible region)

5 Validation Results

5.1 Elimination Consistency

We verified that our estimated fan votes produce correct eliminations:

Method	Correct	Total Weeks	Accuracy
Rank-Based (S1–2, S28–34)	56	56	100.0%
Percent-Based (S3–27)	193	193	100.0%
Overall	**249**	**249**	**100.0%**

Table 2: Elimination consistency validation results

5.2 Sensitivity Analysis

We tested model robustness by adding noise to judge scores:

Noise Level	Rank Method Stability	Percent Method Stability
5%	99.1%	67.3%
10%	98.2%	41.4%
15%	96.8%	28.7%

Table 3: Stability under judge score perturbation

Key Insight: The rank-based method is more robust to score variations because small score changes rarely alter rankings, whereas percentage changes propagate linearly.

6 Controversial Cases Analysis

6.1 Overview

We analyzed four historically controversial contestants where fan votes seemingly overruled judge scores:

Contestant	Season	Placement	Est. Fan Vote %	Advantage
Jerry Rice	2	2nd	20.3%	+4.9%
Billy Ray Cyrus	4	5th	13.6%	+0.2%
Bristol Palin	11	3rd	15.2%	+1.8%
Bobby Bones	27	1st	11.6%	+0.9%

Table 4: Controversial contestants and their estimated fan vote advantages

6.2 Bristol Palin: Most Extreme Case

Bristol Palin (Season 11) represents the most extreme fan-judge disconnect:

- Had the **lowest judge scores 12 times** throughout the season
- Still finished **3rd place**
- Our model estimates she had the **#1 average fan votes** in Season 11 (15.2%)

This validates our model: the only way to explain her survival is through extraordinarily high fan support.

6.3 Bobby Bones: Voting Method Impact

Bobby Bones (Season 27) won despite consistently low judge scores:

- Season 27 used the **rank-based method**
- In rank-based voting, even a 1st-place fan rank provides significant advantage
- As a radio host reaching 150+ stations, he had a built-in voting army

This highlights how the **choice of voting method significantly impacts outcomes**.

7 Implementation Details

7.1 Software Stack

- **CVXPY 1.8.0:** Convex optimization modeling
- **ECOS/SCS:** Interior-point solvers

- **SciPy 1.17.0:** `rankdata` with `method='average'`
- **NumPy/Pandas:** Data manipulation
- **Matplotlib/Seaborn:** Visualization

7.2 Computational Performance

- **Total runtime:** ~2 minutes for all 34 seasons
- **Success rate:** 99.2% (2,067/2,084 contestant-weeks)
- **Optimization convergence:** Typically 10–50 iterations per week

8 Future Work: Ideas for Objectives 2–5

8.1 Objective 2: Rank vs. Percent Method Comparison

Questions to Address:

1. Apply *both* methods to *all* seasons — how often do outcomes differ?
2. Does one method favor fan votes more than the other?
3. Simulate the "judges choose from bottom two" rule (S28+ modification)

Proposed Approach:

- For each season/week, compute eliminations under both methods
- Create a "counterfactual history" — who would have been eliminated differently?
- Quantify fan vote leverage: define a metric $\lambda = \frac{\partial \text{elimination}}{\partial \text{fan vote}}$

8.2 Objective 3: Pro Dancer and Celebrity Characteristics Analysis

Questions to Address:

1. How much does the professional partner impact performance?
2. Do celebrity characteristics (age, industry, homestate) predict success?
3. Are effects different for judge scores vs. fan votes?

Proposed Approach:

- **Mixed-effects regression:**

$$\text{Score}_{ij} = \beta_0 + \beta_1 \cdot \text{Age}_i + \beta_2 \cdot \text{Industry}_i + u_j + \epsilon_{ij} \tag{15}$$

  where u_j is a random effect for pro dancer j

- **Causal inference:** Use propensity score matching to estimate pro dancer effect
- **Feature importance:** Random forest or gradient boosting to rank predictors

8.3 Objective 4: Alternative Voting System Design

Goal: Propose a "fairer" or "more exciting" voting system
Ideas to Explore:

1. **Weighted combination:**

$$S_i = \alpha \cdot P_i^J + (1 - \alpha) \cdot P_i^F \tag{16}$$

   Optimize α to balance skill recognition vs. fan engagement

2. **Momentum-adjusted scoring:**

$$S_i^{(w)} = P_i^J + P_i^F + \gamma \cdot \Delta_i^{(w)} \tag{17}$$

   where $\Delta_i^{(w)} = \text{Score}_i^{(w)} - \text{Score}_i^{(w-1)}$ rewards improvement

3. **Elimination protection:** Top performer each week is immune
4. **Ranked-choice voting:** Fans rank top 3, instant runoff
5. **Skill-tiered voting:** Fan votes weighted by judge score bracket

8.4 Objective 5: Producer Memo

Key Recommendations to Include:

1. Rank vs. percent method tradeoffs
2. When to use "judges choose from bottom two"
3. Balancing competition integrity with entertainment value
4. Managing controversial outcomes

9 Appendix: Data Summary

Statistic	Value
Total contestants	421
Seasons covered	1–34
Fan vote estimates generated	2,067
Unique pro dancers	42
Weeks with data	11 (maximum per season)
Judge score columns	44

Table 5: Dataset summary statistics

9.1 Voting Method by Season

$$\text{Method}(s) = \begin{cases} \text{Rank-based} & \text{if } s \in \{1, 2\} \cup \{28, 29, \ldots, 34\} \\ \text{Percent-based} & \text{if } s \in \{3, 4, \ldots, 27\} \end{cases} \tag{18}$$

10 References

1. COMAP MCM 2026 Problem C: "Data with the Stars"

2. Diamond, S., & Boyd, S. (2016). CVXPY: A Python-embedded modeling language for convex optimization. *Journal of Machine Learning Research*.

3. SciPy documentation: `scipy.stats.rankdata`