

Louis Romeo
CSC 496
HW3 Analysis
10/2/2024

Bunt Analysis for MLB 2023 Season

- **The Original Analysis**

In the first part, we analyze bunts with a runner on 1st and 0 outs using the example CSV file (example_bunt_data.csv), which models the format of the 2430 csv files. The first analysis output will be the following, one variable per line: num-BuntsLaidDown, numUnsuccessfulSacrifices, numSuccessfulSacrifices, numBuntsEveryoneSafe, num-MissedBunts, sacrificeRate, hitRate, failureRate.

The example data includes 6 real bunt attempts to demonstrate accuracy:

- **Missed Bunts (X):** 2 fouled bunts
- **Unsuccessful Sacrifices (Y):** 1 attempt where the batter was safe, but the runner was out at second
- **Successful Sacrifices (Z):** 2 successful sacrifice bunts where the runner advanced, and the batter was out
- **Bunt Hits (W):** 1 play where both the batter and runner were safe

From this data, the probabilities are:

- $X = 2/6 = 0.33$
- $Y = 1/6 = 0.17$
- $Z = 2/6 = 0.33$
- $W = 1/6 = 0.17$

The probability of scoring at least one run from these bunts depends on the advancement of the runner. In this case: **Successful Sacrifices (Z):** The runner advanced, increasing the likelihood of scoring. **Bunt Hits (W):** Both the batter and runner were safe, further enhancing the chances of scoring.

The scenarios where the runner advances (Z and W) combined offer a 50% chance ($Z+W=0.33+0.17=0.50$) of increasing the probability of scoring at least one run.

- **Modification setting X to zero and W determined by the hit rate on bunts with nobody on ($X=0$)**

In this part, we assume no missed bunts ($X=0$) and recalculate the probabilities. We estimate variable W based on the hit rate for bunts meant as base hits, setting $W=0.30$. The remaining probabilities Y and Z are scaled proportionally to maintain $Y+Z+W=1$

- $Y = 0.23$
- $Z = 0.47$
- $W = 0.30$
- $X = 0$

In this modified scenario, the probability of scoring at least one run increases due to a higher likelihood of advancing the runner ($Z+W=0.47+0.30=0.77$). This analysis reflects a significant improvement in run-scoring potential by eliminating missed bunts.

- **Modification that returns X to what it was in the original analysis**

In this part, we return X to its original value (ex. $X=0.33$) while keeping $W=0.30$. The remaining probabilities Y and Z are adjusted accordingly:

- $Y=0.12$
- $Z=0.25$
- $W=0.30$
- $X=0.33$

- **Written Analysis for the probability of scoring at least 1 run**

With a runner on 2nd, the probability of scoring at least one run increases when the bunt is successful or results in a hit. Based on the pocket guide probabilities and assuming typical advancement rates, the combined $Z+WZ + WZ+W$ gives a high likelihood (about 55-77%, depending on the inning and context) of scoring at least one run, particularly in late-game situations where bunting is a strategic move.

This analysis provides insights into the effectiveness of bunting strategies in MLB games with a runner on 1st and 0 outs, using sample data to calculate key probabilities. The modifications where $X=0$ and adjusting W provide the most realistic views on the potential outcomes of bunting.

I used a short example csv file that should be formatted the same as the 2430 test files in the directory. I modified the code from hw2 to download the bunt data and will submit two of the example csv files I used for analysis. Please let me know if there is an issue testing the code or the directory.

