Lab 3

Due on 03/08/2024 at 11:59 pm

Question 1 In class we computed the run expectancy matrix for the 2016 season. We used this quantity to assess the value of stolen bases and we computed the marginal break even stolen base percentage required to justify an attempt. Do the following:

- Rank catchers by their ability to prevent runs via catching runners attempting to steal bases. Report a top 10 list for most effective catchers at preventing runs via the stolen base. Also report the worse catchers.
- Compute break even stolen base percentages at different base out states and for different innings. You can consider a minimum attempt threshold. When are good times to attempt or not a stolen base?

Question 2 In the Simulation Notes we considered team specific transition probabilities for one base out state corresponding to the St. Louis Cardinals in 2016. Do the following:

- Build the entire transition probability matrix for the St. Louis Cardinals. This matrix should be constructed using the normalized versions of the reliable probabilities with $K = 1274$. Hint: you may want to use the code in the notes within a loop, and then convert from long format to wide format using something like

```
library("reshape2")
dcast(melt(foo, id.vars=c("STATE", "NEW.STATE")), STATE~NEW.STATE) %>%
    replace(is.na(.), 0)
```

- Display the typical states after 3 PAs to start a half-inning for the 2016 St. Louis Cardinals. Were the 2016 Cardinals expected to end a half-inning more frequently than the average team after 3 PAs?
- Use the `simulate_half_inning` function and the Runs matrix `R` in the notes to construct a 2016 Cardinals specific RE24 matrix.

Question 3 Problem 5 on page 135 of Analyzing Baseball Data with R

Question 4 Problem 1 on page 226 of Analyzing Baseball Data with R

Question 5 Problem 3 on page 227 of Analyzing Baseball Data with R