Mallory Folliard and Ella L'Esperance
CS152: Web Scraping Lab

In this analysis, our goal is to apply Bayes' Theorem to explore the relationship between having a night game and the probability of winning in professional football, specifically zooming into the New England Patriots' performance in the past 5 seasons. Bayes' Theorem provides a statistical method for updating the probability of a hypothesis as more evidence becomes available. Our hypothesis (Event A) is that a game is a night game, and our evidence (Event B) is the team winning a game.

Defining Our Events:

- Event A (Hypothesis): The game is a night game, defined as starting after 5:30 PM. This percentage is the proportion of night games to total games in the 2019-2023 seasons. This probability was found by determining the ratio between all of the night games in that time frame by all of the games played.
- Event B (Evidence): The team wins the game. This is based on the games won in the 2019 to 2023 season compared to total games in that same time frame. This probability was found by dividing total wins by total games.
- Event B given A: We found this data by calculating the number of games won that were night games and dividing that by total number of games in those past 5 seasons.

Data Utilization

The data for this analysis was scraped from the Pro Football Reference website, specifically the New England Patriots' Franchise website as well as the 2019 - 2023 schedule sites. We scraped the game times from the 2023 season table and overall outcomes (win/loss) for each game. This dataset provided us with the necessary information to calculate the prior probability P(A) and the likelihood P(B), and P(B|A). The total number of games, the number of night games, and the number of wins were all counted to support these calculations. In the end, our goal was to find P(A|B), this being the probability of winning given the game is a night game.

Actionability of the Analysis

The result of this analysis could be actionable in strategic planning for the team. For instance, if the final probability indicates a significantly higher chance of winning night games, the team might advocate for more night games in future schedules. However, as is true with our result being .24 chance of winning if the game is a night game, the data does not significantly impact the winning probability, and the team might focus on other areas for improvement rather than game timing.

Conclusion

We believe this analysis demonstrates a practical application of Bayes' Theorem in sports analytics applied to the Patriots 2019 to 2023 seasons. While the current analysis is limited to one team and only five seasons, it could be expanded to include multiple teams and more seasons for a more comprehensive understanding of the impact of game timing on performance in professional football. We are also unsure as to how much impact a team can have on their game timing. Overall we felt as though our python code demonstrates understanding of scraping as well as applying Bayes theorem on scraped data on the past five NE patriots seasons.