Free Throw Shooting in College Basketball: Is There Really a "Home Court Advantage?"

Abstract: This project examines regular season free throw shooting percentage for every college basketball team in the years 2003-2019. A Bayesian approach relying upon the beta and binomial distributions was used to assess this research question. After accessing data from the Kaggle website, analysis was conducted to determine whether the home teams in a stadium have a free throw shooting advantage over the away teams. The prior and posterior distributions from home and away teams were found. Although there were a wide range of free throw shooting differences between home and away teams throughout all college basketball stadiums nationwide, only a slight overall difference was found in free throw shooting percentage between home and away teams.

Introduction

The research question we decided to test was, "do college basketball teams playing on their home floor shoot better from the free throw line than the visiting team?" The value in conducting this research was to identify how much of an effect stadium crowds have on distracting the opposing team's free throw shooters. Each crowd in the country does their best to employ all sorts of frenzied tactics to help their team gain an advantage at the free throw line each game, but is their influence really even recognized? We performed the test using data from the 2003-2019 NCAA college basketball seasons.

Literature Review

After beginning our project, we found a *New York Times* article entitled "Free Throw Distraction: The Best Fans in the N.C.A.A.". This article described a study that was conducted to identify which home crowds affect opposing team free throw shooters the most. The authors found that Arizona State University, with their implementation of the "Curtain of Distraction", had the greatest effect on opposing teams' free throw shooting. See reference 1 to find the citation of this article.

Methods

The data were obtained on Kaggle's website. Kaggle held a competition for predicting the college basketball tournament and the competition creators uploaded various data sets from past game results. We used the "Regular Season Detailed Results" file for our calculations since it provided all the statistics for each team in every regular season game dating back to the 2003 season. See reference 2 to find the citation for the data on the Kaggle website.

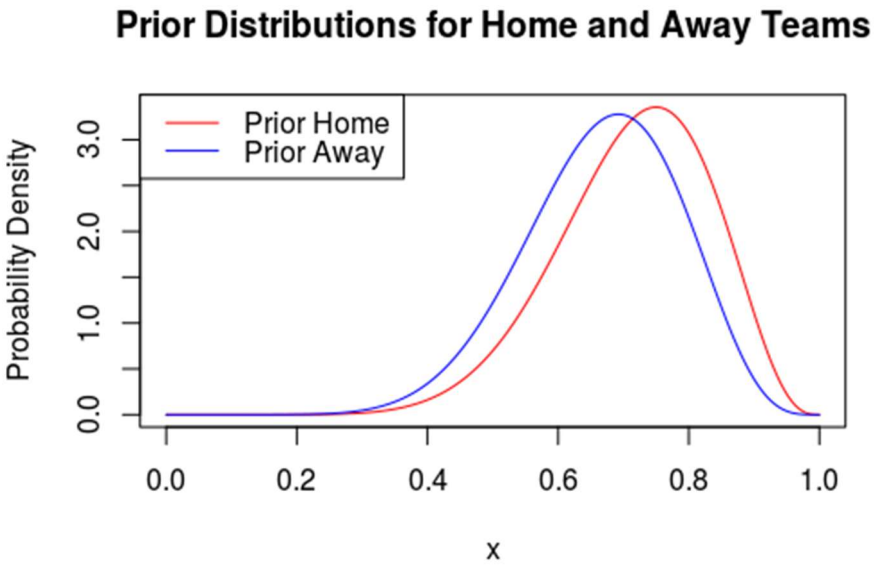
As our model of choice, we picked a beta-binomial, which requires the following assumptions:

- The observed data has a fixed number of trials
- Each trial has a constant probability of success
- Each trial outcome is independent of the other outcomes

The first assumption is met because a free throw must be completed and thus there are no unfinished trials. The second assumption is met because we assume that each shooter's probability of making the free throw does not change over time. We are not entirely sure that the third assumption is met, because free throws may have a different level of importance throughout each game which could affect the shooter's ability to make the free throw.

We tested our data by treating each college basketball stadium as a unit while looking at the home team and away team in each unit. Then, y_k (number of free throws made) is independently and identically distributed by a binomial distribution for a home team with parameters n_k (number of free throws shot) and θ_{home} (probability of a made free throw). Likewise, x_k (number of free throws made) is independently and identically distributed by a binomial distribution for a home team with parameters t_k (number of free throws shot) and θ_{away} (probability of a made free throw). The subscript $k = 1, \ldots, 351$ because there are 351 college basketball stadiums in the nation.

Our prior distribution is a beta distribution with parameters a and b. For our prior distribution for home teams, we chose a = 10 and b = 4. For our prior distribution for away teams, we chose a = 10 and b = 5. We wanted the means for both home and away teams to be around 0.7, and given our prior values, the mean for home teams is 0.714 and the mean for away teams is 0.667. We originally had chosen larger values for a and b as our prior values. However, the variances seemed too small to represent our uncertainty, so we made our prior values smaller. Our prior variances with our chosen a and b values are around 0.014.

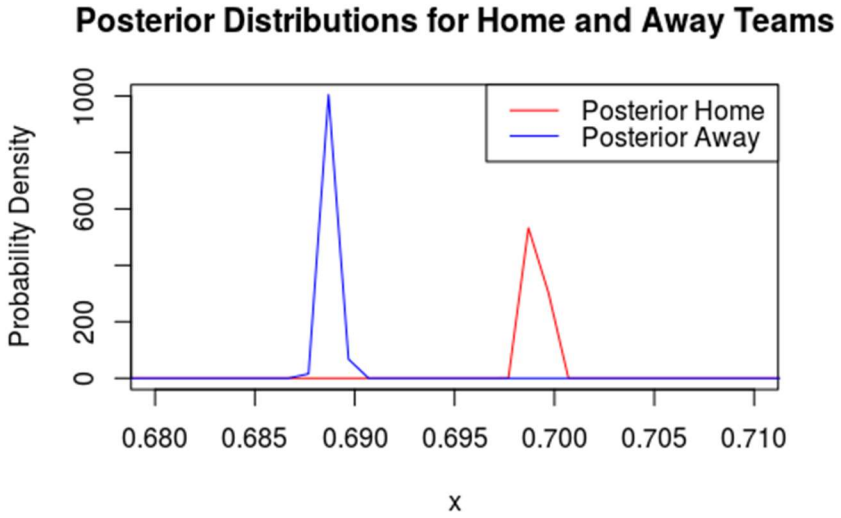


After finding the posterior distribution for home and away teams in each stadium, we will find the posterior mean of the differences between home and away teams in each stadium to answer the research question.

Results

Total free throws made for home teams: 1,202,948
Total free throws attempted for home teams: 1,720,635
Free throw percentage for home teams: 69.913%
Total free throws made for away teams: 1,013,900
Total free throws attempted for away teams: 1,471,988
Free throw percentage for away teams: 68.879%
Total number of games played:
78,584
Total number of seasons used: 17

After performing posterior inference, we found our posterior values for a and b to be as follows:
$a^*_{home} = 1202958$
$b^*_{home} = 517691$
$a^*_{away} = 1013910$
$b^*_{away} = 458093$



Since the probability density for the posterior distributions is much greater than the prior distributions, the prior distributions are basically a horizontal line at 0 compared to these posterior distributions. Thus, we did not include a second graph.

Credible interval for home teams: (0.6984, 0.6998)

- We can predict that the free throw percentage for home teams in college basketball stadiums is between the interval of (0.6984, 0.6998) with 95% posterior probability.

Credible interval for away teams: (0.6880, 0.6895)

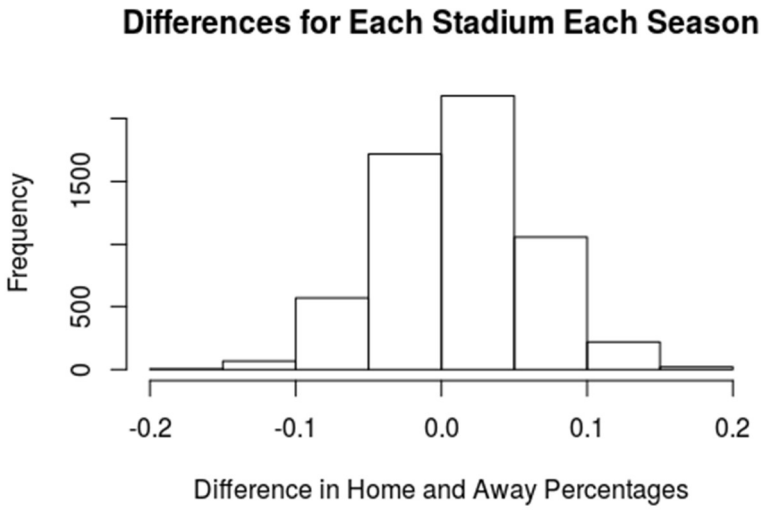
- We can predict that the free throw percentage for away teams in college basketball stadiums is between the interval of (0.6880, 0.6895) with 95% posterior probability.

Our posterior mean for the home team free throw percentage is 0.6991 and our posterior mean for the away team free throw percentage is 0.6888. Thus, the difference in the two means is 0.0103. Our posterior variance for the home team free throw percentage is 1.222×10^{-7} and our posterior variance for the away team free throw percentage is 1.456×10^{-7}. Thus, the ratio of our posterior variances is 0.8395.

Credible interval for difference in means: (0.0093, 0.0113)

- The difference in means of free throw percentage between home teams and away teams in college basketball stadiums is in the interval (0.0093, 0.0113) with 95% posterior probability.

The graphic to the right shows a histogram of the mean difference in home team and away team free throw percentages for each stadium and each season between 2003 and 2019. We can see that most stadiums see very little difference in free throw shooting percentage between home and away teams.



Discussion

Based on our posterior results and the graphs above, we discovered that there is a difference in free throw percentages between the home team and the away team in college basketball stadiums. If we had more time, we could better answer our question by adding the variable of attendance per game for each of the games within the data. Obviously, each game does not have the same total of attendees, and average school attendance can range anywhere from over 20,000 people to as few as 1,000. Follow-up research could look into the effect that the attendance in each stadium has on free throw percentages for home and away teams. This could help us begin to calculate and place a value on the importance of a game being played at home for a certain team or the value of having an extra thousand attendees to a game in a certain stadium.

We did not adjust for in-game factors. These factors could be ideas such as which player was shooting the free throws for a team (whether the majority of the free throws were shot by "bad" or "good" free throw shooters) or the game time when the free throws were taken (beginning of game vs. end of game may have a different level of importance and pressure for the free throw). We hope to look into further research to identify how these factors and others could affect free throw shooting.

References

1) Quealy, Kevin, and Justin Wolfers. “Free-Throw Distraction: The Best Fans in the N.C.A.A.” *Nytimes*, New York Times, 12 Mar. 2015, www.google.com/amp/s/www.nytimes.com/2015/03/13/upshot/free-throw-distraction-the-best-fans-in-the-ncaa.amp.html.

2) “Google Cloud & NCAA® ML Competition 2019-Men's.” *Kaggle*, Kaggle, Mar. 2019, www.kaggle.com/c/mens-machine-learning-competition-2019/data.