Programming Assignment 3

Anthony Gallante
MSDS-452-DL
Graphical, Network, and Causal Models
02 21, 2026

Abstract

In this assignment, we investigate the causal effect that playing away has on college basketball free throw attempts. We begin by building a simple directed acyclic graph (DAG) using python's pgmpy library before validating its assumptions using the DoWhy library. Using 45,324 player performances in the 2025-2026 season, we estimate that visiting players shoot approximately two percentage points lower than they do at home.

Introduction

Free throws in basketball are a unique event in sports, in that an athlete is given an uncontested chance to score. The player who is awarded a free throw attempt is always positioned 15 feet directly in front of the basket while opponent athletes stand aside until the ball is thrown. Rule 8, Section 5 Article 1.f prohibits opponent players from interfering with this attempt, stating: "No opponent (player or bench personnel) shall disconcert (e.g., taunt, bait, gesture or delay) the free-thrower" (National Collegiate Athletic Association 2023). Spectators, on the other hand, are not subject to this rule. In fact, it has become a tradition for home-team spectators to be as disruptive as possible when a player from the opposing team steps to the line. An Arizona State University student fan group famously operates the "Curtain of Distraction" during free throws, where a physical curtain is opened to reveal students—and in one case, USA Olympic champion Michael Phelps—wearing costumes in an attempt to distract opponents (Gharib 2024).

In this assignment, we build and validate one plausible causal model that can be used to estimate the disadvantage of shooting free throws as the visiting team, if there is one.

Literature Review

An Analysis of NBA Home Court Advantage (Fettig et al. 2024)

This paper suggests that larger crowds are linked with better field goal shooting efficiency and higher scoring for the home team, but not win rate. Additionally, while referees tend to have a slight bias toward the home team, the researchers did not find that crowd size does not seem to further exaggerate this bias. Free throw performance specifically was not observed to change with crowd size when the referee bias was taken into consideration.

Quantifying Home Court Advantages for NCAA Basketball Statistics

(Bommel et al. 2021)

Once again, researchers report a referee bias in favor of the home team, but van Bommel reports that crowd size could amplify the bias at the college level. University teams that perform at higher levels were also observed to benefit more from these biased home court benefits. This paper also shares the

Influence of Home-Court Advantage in Elite Basketball (Mochales Cuesta et al. 2024)

Cuesta et. al. suggest that travel distance and eastward jet lag affect visiting teams much more than crowd interactions. The researchers do acknowledge that large crowds are more likely to affect free throw shooting in high pressure situations.

Methods

Data

The dataset of 45,324 player performances was used for this analysis, containing a combination of team and player level records made available through the HoopR R package (source). A data dictionary of the relevant data fields is shown below.

Variable	Data Type	Description
minutes	int	The number of minutes that a player has played
home_away	bool	0 if this is a home game. 1 if this is an away game.
opponent_fouls	int	The number of fouls committed by the opposing team
free_throws_attempted	int	The number of free throws attempted
free_throws_made	int	The number of free throws made
free_throw_pct	float	The percentage of free throws made

Table 1: Independencies found using pgmpy.

Causal Assumptions

One possible directed acyclic graph (DAG) that could be used to model this system is shown below in Figure 1. Free throw attempts are dependent on both the amount of time a player spends on the court as well as the number of fouls committed by the opposing team. Here, the assertion is made that the number of successful free throws is somehow influenced by the player being at a different home court (the treatment variable). Based on the literature referenced above, we should see little to no effect, but since we have not conditioned on crowds with distracting behavior, we might

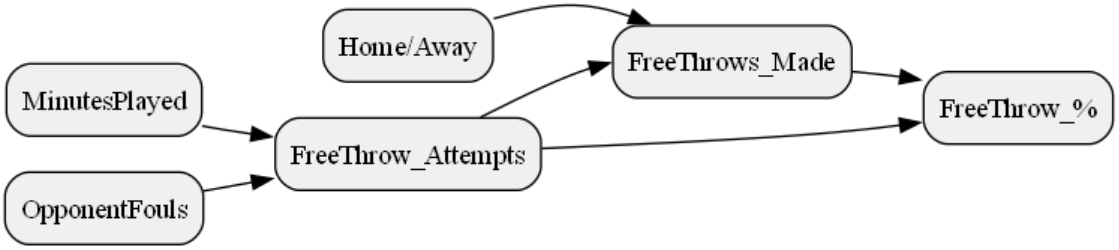

Figure 1: A possible DAG that describes the causal interactions in a free throw attempt

Additionally, the relationship here could be further confounded by a "Pressure" variable which might be defined by some relationship between the difference in score (positive if winning, negative if losing) and the time remaining in the game. It is also worth noting that defensive aggression and the number of intentional fouls rises significantly in the final minutes of a close basketball game as teams attempt to force clock stoppages and turnovers. Data at this level is not readily available.

Results

The DAG above implies several testable independencies, shown below in Table 2.

Testable Independencies
home_away ⊥ minutes
free_throw_pct ⊥ minutes | free_throws_attempted
opponent_fouls ⊥ free_throw_pct | free_throws_attempted
free_throw_pct ⊥ home_away | free_throws_made, free_throws_attempted
minutes ⊥ free_throws_made | free_throws_attempted
home_away ⊥ free_throws_attempted
opponent_fouls ⊥ free_throws_made | free_throws_attempted
opponent_fouls ⊥ minutes
opponent_fouls ⊥ home_away

Table 2: Independencies found using pgmpy.

In his book, Ness shares Python code using the pgmpy library that uses chi squared tests to test for statistical independence (Ness 2025), but these will certainly fail with over 45,000 samples.

The DoWhy Python library makes several other validation methods available to the analyst (Sharma and Kiciman 2020). In this assignment we compare the measured effect with that of a placebo and to similar treatments that may be confounded by an unknown random variable.

Conclusions

Using DoWhy to compute both linear and nonlinear causal effects, we estimate that players on visiting teams perform about 2 percentage points lower than they would at home games. The reader should be reminded that the dataset used does not include crowd distractions specifically, so perhaps the results are underwhelming. However, a measurable effect was calculated with this relatively small dataset and simple DAG.

Interestingly enough, when we look at the 111 player performances played at Arizona State University, who has a reputation for distracting their opponents, we observe a greater

impact to free throw shooting. The average effect doubles to 4% and the conditional average treatment effect range is expanded significantly, with some players shooting 22 percentage points worse than otherwise, with some shooting up to 26% better. To gain a better understanding of the "Curtain of Distraction," however, we would likely want to compare player free throw shooting performance to include many seasons, both before and after 2013, when the curtain was unveiled for the first time.

Bibliography

[1] National Collegiate Athletic Association, “2023–24 NCAA Men's Basketball Rules.” Indianapolis, IN, Aug. 2023. [Online]. Available: https://www.ncaa.org/playingrules

[2] A. Gharib, “Top men's college basketball free throw distractions of the season.” [Online]. Available: https://www.espn.com/mens-college-basketball/story/_/id/39748477/college-basketball-free-throw-distractions-2023-24

[3] S. Fettig, M. Guldan, M. Law, and Y. Ritov, “An analysis of NBA home court advantage,” 2024.

[4] M. van Bommel, L. Bornn, P. Chow-White, and C. Gao, “Home sweet home: Quantifying home court advantages for NCAA basketball statistics,” *Journal of Sports Analytics*, vol. 7, pp. 25–36, 2021, doi: 10.3233/JSA-200450.

[5] I. Mochales Cuesta, S. L. Jiménez-Sáiz, A. L. Kelly, and Á. Bustamante-Sánchez, “The influence of home-court advantage in elite basketball: A systematic review,” *Journal of Functional Morphology and Kinesiology*, vol. 9, no. 4, p. 192, 2024, doi: 10.3390/jfmk9040192.

[6] R. O. Ness, *Causal AI*. Shelter Island, NY: Manning Publications, 2025.

[7] A. Sharma and E. Kiciman, “DoWhy: An End-to-End Library for Causal Inference,” *arXiv preprint arXiv:2011.04216*, 2020, [Online]. Available: https://arxiv.org/abs/2011.04216