Personalized Foul Trouble

An Analysis of Fouling Out in the NBA

Sacramento Kings Case Competition Winners

Dani Chu, Lucas Wu, Matthew Reyers & Conor Doyle

Simon Fraser University

Contact Information:

Dani Chu
dani_chu@sfu.ca
danichusfu.github.io

Matthew Reyers
matthew_reyers@sfu.ca

Lucas Wu
yifan_wu@sfu.ca

Conor Doyle
conor_doyle@sfu.ca

Abstract

NBA players foul out when they are called for 6 fouls in one game. Coaches in the NBA usually say that a player is at risk of fouling out of a game when they have 1 more foul than the current quarter. This is called the Q+1 rule. This generic foul trouble calculation treats all players equally, despite many players having known tendencies with respect to foul acquisition. Our objective is to creatively outline objective and personalized guidelines for a coaching staff to handle in-game foul management. To do so we used NBA play by play data from the 2014-2015 season to the 2016-2017 season and looked specifically at the fouling events for each player. We then used a Cox Proportional Hazards Regression to relate predictors to survival time. We then fit personalized survival curves for each player and foul number and estimated a rate parameter from a truncated exponential distribution that would best fit that curve. Finally, we used those curves to simulate the amount of time that a player would have remaining in a specific game, given the number of fouls they have acquired. The results of our estimations were then built into an interactive tool for the Sacramento Kings roster in 2016-2017. The tool can be expanded to any roster and year as required. We find that this personalized and data driven method to understanding when a player will foul out is more useful than a catchall approach.

Introduction

Current Practice

- Managing the minutes of a team's star player and maximizing their output is a difficult task
- Currently coaches typically use the "Q+1" methodology detailed in Maymin and Shen's research on early foul trouble [4] to determine if a player is in foul trouble
- If a player has more fouls than the current quarter of play they should be benched (ex. 3 fouls in the 2nd quarter.)

Shortcomings and Counterexample

In Ben Falk's article "The Trouble with Foul Trouble" [2] he recounts a scenario that unfolded in the first round of the 2018 NBA playoffs between the Cleveland Cavaliers and the Indiana Pacers.

- Victor Olatipo, the sole all-star of the Pacers picked up 2 fouls in the first quarter and an additional foul in the second quarter over 8 minutes of total playing time
- Oladipo would go on to play 20 minutes in the second half without picking up another foul and the Indiana Pacer lost the game by 3 points
- Coach Nate McMillan sacrificed Oladipo's minutes in the first half in order to maintain flexibility at the end of the game
- Falk points out that no foul management strategy can save a player's minutes, it can only limit them and that only changes in strategy and role can have an affect on saving a player's minutes
- "A player will play less in exchange for the ability to control when he plays."

Previous Work

- Katherine Evans suggests using a survival model to estimate the time until a player commits his next foul
- Evans considers how emotions affect the rate at which players pick up fouls, which she calls "tilting".
- Evan's shows that fouling rates differ depending on the number of fouls the player has

Main Objectives

- Creatively outline objective guidelines to present to a coaching staff about how to handle in-game foul management.
- Suggest an improvement to the current "Q+1" foul management strategy identified by Maymin
- Present personalized strategy at a player level for foul management
- Use Evan's process of fitting survival curves to different foul levels and extend it by fitting exponential distributions to the curves and combining them to gather a time to foul out distribution instead of simply a time to next foul distribution.
- Take into account Falk's suggestions and identify foul trouble on a player by player basis, are more lenient earlier on, and help coaches rearrange minutes instead of limiting them.
- Providing the coaching staff with better information than they used to.
- Trust the coaching staff to make a more informed decision regarding foul trouble

First Model: Cox Proportional Hazards

Survival analysis is an application of statistical methods focused on understanding the time until the occurrence of an event. Often used for death and customer churn calculations, survival analysis can be logically applied to fouling out in basketball. A common component of survival analysis problems is that of censored data. Data is censored when a measurement is only partially known, often due to the end of testing prior to the death of a given unit. Dealing with censored data requires different methods than basic non-censored data. One of the well tested models for this case is then the Cox Proportional Hazards Regression which is used to relate predictors or covariates to survival time. The effect in this model in the context of basketball is then the probability of picking up the next foul which represents the risk of failure. We considered three covariates in our model: player position (Guard, Forward, Big), number of fouls already committed this game, and an indicator variable for 3+ years of experience.

$$\lambda(t|Z) = \lambda_0(t)\exp(\beta_1{\cdot}\text{Position}+\beta_2{\cdot}\text{Fouls}+\beta_3{\cdot}I(\text{Seasons in League} \geq 3))$$

Results of this approach serve primarily as a guideline due to the underlying assumption of constant hazard rate required by the model. Our data review found this not to be the case in our data set for some of our variables, violating one of the assumptions of the Cox Proportional Hazard model.

Second Model: Survival Curves

The notion of survival curves can be applied to the problem of foul trouble by treating each foul as a death. A player is deemed to have survived at a given foul level if, at game's end, they have not obtained another foul. Following Evan's approach, we can then create survival curves for each player separated by a given foul number to incorporate "tilt" [1]. Though these curves are stepwise functions, they bare close resemblance to the exponential distribution (see Figure 1). We use this similarity to justify the fitting of exponential distributions to each set of curves, as done previously by Evan's [1].

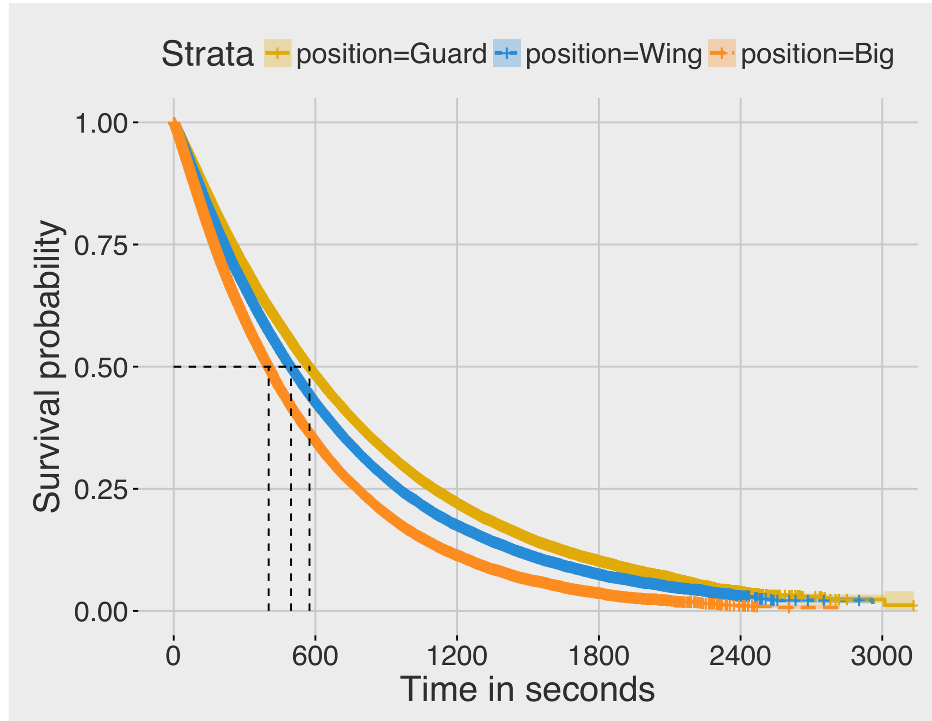


Figure 1: Rudy Gay's Foul Rate by current number of fouls from 2014-2016.

The algorithm for our approach then becomes:

1. Fit appropriate exponential distributions to each curve through numerical integration
2. Generate 5000 samples for each player from each of their foul level exponential distributions
3. Calculate time to foul out for each of these simulations from foul j for player i, $\sum_{f=j}^{5} t_{i,f}$
4. Bootstrap samples of size 1000 to estimate quantiles of the foul out distribution

The exponential distributions used are truncated at the maximum playing time a player has spent at a given foul level in the training set (2014-2016). These are to be updated as players continue to surpass these values and add more data into the analysis. The results of this model are then passed to the applet for coach and front office accessibility.

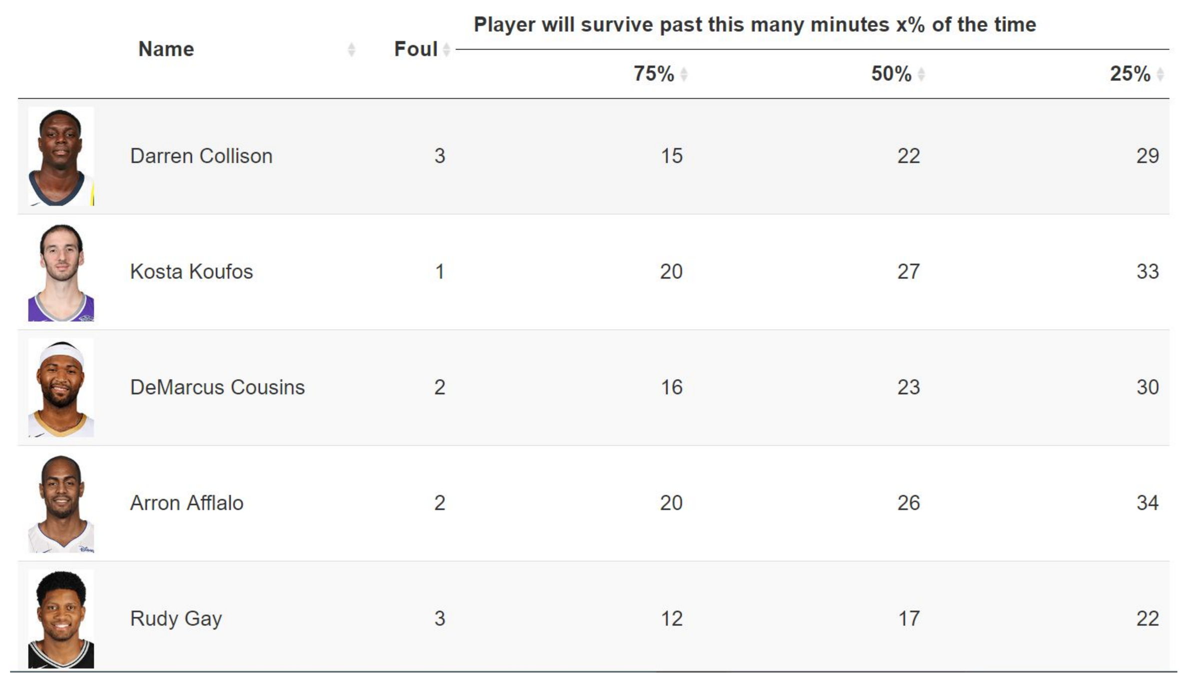
Model Improvements

We would like to continue our work and extend our estimation of time to foul out through a Bayesian method. The foul rate (parameter λ) at each foul level would be informed by a prior distribution of foul rates from players of the same position or defensive style and the likelihood function for the observed data at each foul level for each player. This would give us more stable estimates for those with small amounts of data. The defensive style categorization would be done through a clustering of a player's defensive tendencies. These tendencies could be defined by common foul locations, foul types, defensive matchups and other defensive statistics.

Applet Interaction

For practical in game use of the survival curve approach, an applet is required as the calculations are not as trivial as " Q + 1". To interface with our Time to Foul Out tool, an Assistant Coach can simply inputs the number of fouls each player has and is fed back the estimates on game time remaining. The Assistant Coach can then communicate the information to the Head Coach when necessary.

By providing sample quantile estimates, coaches can decide how aggressively or conservatively they want to manage a player's expected remaining play time. An example of the applet. using the Sacramento Kings is provided in Figure 2.



Name	Foul	Player will survive past this many minutes x% of the time		
		75%	50%	25%
Darren Collison	3	15	22	29
Kosta Koufos	1	20	27	33
DeMarcus Cousins	2	16	23	30
Arron Afflalo	2	20	26	34
Rudy Gay	3	12	17	22

Figure 2: Foul Trouble Tool. Can be used live in games to update the time to foul out after a player picks up a foul. Used here on the 16-17 roster of the Sacramento Kings.

Conclusions

Fouls are a long standing component of basketball that is governed by an archaic coaching rule dictating when a player should sit or play. Researchers close to the world of basketball have put forth alternatives such as survival curves, which we have worked to extend. By adding on a simulation component to generate playing time bounds at given quantiles, coaches now have better access to informed decision making. With the creation of the applet, all of this information is now available in the same time it would take to calculate the old "Q + 1" rule.

References

[1] Evans, K. NBA Fouls - Survival Analysis. `https://causalkathy.com/2017/02/21/nba-fouls-survival-analysis/`, 2017. [Online; accessed April-2018].

[2] Falk, B. The Trouble with Foul Trouble. `https://cleaningtheglass.com/the-trouble-with-foul-trouble/`, 2018. [Online; accessed April-2018].

[3] Klobuchar, A. NBA Foul-Substitution Analysis . `http://aidanklobuchar.com/home/data-based-writings/fouls/`. [Online; accessed April-2018].

[4] Maymin P. Maymin, A. and Shen E. How much trouble is early foul trouble? Strategically idling resources in the NBA. *International Journal of Sport Finance*, 7(4):324–339, 11 2012.

[5] Rosenmai, P. Fitting an Exponential Curve to a Stepwise Survival Curve. `https://eurekastatistics.com/fitting-an-exponential-curve-to-a-stepwise-survival-curve/`, 2016. [Online; accessed April-2018].

[6] Saini, U. & Evans, K. Extended NBA Play By Play Data. `http://eighthirtyfour.com/data`, 2017. [Online; accessed April-2018].

Acknowledgements

A big thank you to Luke Bornn and his team at the Sacramento Kings for providing the opportunity to explore the exciting world of basketball analytics and to Katherine Evans for her inspirational work and for making her data publicly available.