Foul Accumulation in the NBA

Dani Chu and Tim B. Swartz *

Abstract

This paper investigates the fouling time distribution of players in the National Basketball Association. A Bayesian analysis is presented based on the assumption that fouling time distributions follow a gamma distribution. Various insights are obtained including the observation that players accumulate fouls at a rate that increases with the current number of fouls. We demonstrate possible ways to incorporate the fouling time distributions to provide decision support to coaches in the management of playing time.

Keywords : Bayesian analysis, censoring, constraints, failure time distributions, Markov chain Monte Carlo, predictive inference.

*D. Chu is an MSc candidate and T. Swartz is Professor, Department of Statistics and Actuarial Science, Simon Fraser University, 8888 University Drive, Burnaby BC, Canada V5A1S6. Both authors have been partially supported by the Natural Sciences and Engineering Research Council of Canada. The authors thank two anonymous reviewers and the Associate Editor for valuable comments that improved the paper.

1 INTRODUCTION

In the National Basketball Association (NBA), a player fouls out of the game and is ineligible to play after committing their sixth personal foul. Coaches therefore tend to "sit" players (i.e., temporarily remove them from the game) when their fouls accumulate. The intention of the coach is to have key players available later in the game.

A rough guideline that has been followed by NBA coaches is that a player is instructed to sit when their number of fouls reaches $Q + 1$ where $Q = 1, 2, 3$ is the quarter of the match. For example, when a player attains two fouls in the first quarter of a match, the coach will remove the player from the game. Similarly, when a player attains three fouls during the second quarter of a match, the coach will remove the player from the game.

The above coaching tradition does not appear to be informed by data, and does not take into account specific player tendencies. For example, it may be possible that a particular player can continue to be effective and avoid fouling. In sporting practice, there exist traditions that are on the level of folklore, and upon closer inspection, do not appear optimal. For example, in hockey, the tradition had been for a team to pull its goalie when trailing with about one minute remaining in a match. However, it has been suggested through statistical modelling and simulation that goalies should be pulled with approximately three minutes remaining (Beaudoin and Swartz 2010). The new recommendations appear to have made an impact in NHL practice (Davis and Lopez 2015). Another example of a misguided sporting tradition involves the over-reliance on popular baseball statistics such as batting average. As is well known, the Moneyball phenomenon (Lewis 2013) highlighted alternative baseball measures such as on-base percentage which serve as better predictors of success. In the sport of football, Yam and Lopez (2018) use methods of causal inference to assess the impact of punting on fourth down in the National Football League.

In this paper, we use data and statistical modelling to investigate various questions associated with foul accumulation in the NBA. For example, do all players foul at the same rate? Does the distribution of playing time between the $(n-1)$st foul and the nth foul, $n = 1, \ldots, 6$ depend on n? Are there differences in the time-between-fouls (fouling time) distributions according to playing position? With respect to foul accumulation, can we advise coaches when to sit their players?

The topic of NBA substitution patterns is a topic that has been mostly discussed on blog sites. For example, Rochford (2017) use item response theory and Bayesian modelling to draw various insights with respect to NBA fouls. In particular, Rochford (2017) draws attention to the relationship between fouling and salary with the suggestion that higher paid players are treated preferentially with respect to foul calls. Klobuchar (2018) investigates the impact on win shares from the "early" substitution of players due to fouls. Falk (2018) examines playing minutes when a coach employs a foul management strategy. The investigation suggests different strategies, including changing a player's defensive assignment to decrease a player's foul rate and rearranging the player's minutes. In the journal article by Maymin, Maymin and Shen (2012), the impact of early foul trouble is assessed using tools from finance. Of note, Maymin, Maymin and Shen (2012) suggest that teams exhibit poorer performance if they continue to play foul-plagued starters. Evans (2017b) propose a conditional risk set model for ordered events to model a player's time-to-foul while including covariates such as the point differential, and time remaining in the game among others.

In Section 2, we begin with an exploratory data analysis where we assess the conjecture that the fouling time is exponentially distributed. We suggest that the gamma distribution provides a more realistic fouling time distribution. We also investigate the impact of player position and the impact of foul level on the fouling time distribution. In Section 3, we use the gamma distribution to build a stochastic model which incorporates unknown parameters. The model is Bayesian and requires the specification of prior distributions and computational strategies to assess the parameters. The exploratory analysis in Section 2 helps us specify the prior distribution in Section 3. A predictive distribution is then introduced which may be used to address the practical intentions of coaches. The models are implemented on NBA data in Section 4 where interesting insights are obtained with respect to the fouling tendencies of players. We conclude with a short discussion in Section 5.

2 EXPLORATORY DATA ANALYSIS

2.1 The Fouling Time Distribution

It has been suggested that the exponential distribution may be appropriate for the distribution of fouling times (Evans 2017a). Some simple thought experiments reveal that the validity of this assumption is less than clear-cut. For example, if the fouling time is exponentially distributed, then this time between fouls satisfies the *memoryless property*. This implies that a player who has just stepped onto the court has the same probability of fouling within a period of time compared to the situation where the player had been on the court for a period of time. The memoryless property seems suspect as it is well known that a tired basketball player may have difficulty moving his feet into a good defensive stance, and is therefore more likely to commit a foul than a fresh player. On the other hand, it may be argued that a fresh player may be overly excited and aggressive, and may be more prone to foul than a player who has been on the court for a while.

In order to test the suitability of the exponential distribution, we introduce some standard failure time notation. Consider the ith player who has committed his $(n-1)$st foul, and this occurs in a match which we label the jth match. We denote $X_{ij}^{(n)}$ as the time played between the $(n-1)$st and nth foul, $n = 1, \ldots, 6$ and $j = 1, \ldots, m_{in}$ where m_{in} is the number of matches in which the ith player has reached the nth foul level. It is therefore apparent that $m_{i1} \geq m_{i2} \geq \cdots \geq m_{i6}$ since a player must reach all foul levels less than n to reach the nth foul level. It is possible that the time to foul $X_{ij}^{(n)}$ is unobserved and there is a potential censoring time $C_{ij}^{(n)}$. In this case, the corresponding observed dataset for the ith player at the nth *foul level* is given by $(Y_{i1}^{(n)}, \delta_{i1}), \ldots, (Y_{in_m}^{(n)}, \delta_{in_m})$ where $Y_{ij}^{(n)} = \min(X_{ij}^{(n)}, C_{ij}^{(n)})$ and

$$\delta_{ij} = \begin{cases} 0 & X_{ij}^{(n)} \leq C_{ij}^{(n)} \quad \text{(uncensored)} \\ 1 & X_{ij}^{(n)} > C_{ij}^{(n)} \quad \text{(censored)} \end{cases} .$$

In this application, it is important to note that the censoring mechanism involves random right censoring rather than fixed right censoring. Should a player not commit the nth foul,

$n = 1, \ldots, 6$, we can think of the nth foul as randomly censored. In a medical application with fixed right censoring, this corresponds to an experiment which concludes at the same time for all subjects. A detailed treatment of the statistical analysis of failure time data is given by Kalbfleisch and Prentice (2002).

In our investigation of the exponential distribution to model fouling times , we first note that the exponential distribution is a special case of the gamma distribution. Here, we use the parameterization $X \sim \text{Gamma}(\alpha, \beta)$ such that $\text{E}(X) = \alpha/\beta$. We consider alternative data to that used in Section 4 to avoid the perils of "double use of the data." Specifically, we consider data from the 2012/2013 NBA regular season involving players i at the nth foul level who have $m_{in} \geq 30$ observations. This provides 1,010 player-foul combinations involving 376 unique players.

We test the fit of the exponential distribution against the fit of the gamma distribution at the 0.05 level of significance for each of the 1,010 datasets. That is, we carry out likelihood ratio tests of the null hypothesis, $\text{H}_0 : \alpha = 1$ (i.e., exponential) against the alternative hypothesis, $\text{H}_0 : \alpha \neq 1$ (i.e., gamma but not exponential). Expressions for the maximum likelihood estimators of α and β under right random censoring are given by Harter and Moore (1965). These are estimated using the *mle* function from the stats4 R package which finds the minimum of a specified negative log-likelihood using an optimizer.

In the context of multiple comparisons, we reject 20% of the null hypotheses (much more than 5% of null hypotheses that would be expected due to random variation). An associated p-value for the number of rejected tests corresponds to $\text{Prob}(W > 202) \approx 0$ where $W \sim \text{Binomial}(1010, 0.5)$. This provides evidence that the gamma distribution is more appropriate than the exponential distribution. More powerful goodness-of-fit tests such as those based on the empirical distribution function (see D'Agostino and Stephens 1986) would likely result in higher rates of rejection of the null hypothesis.

We further investigate the adequacy of the exponential and gamma distributions by comparing exponential and gamma survival curves (based on maximum likelihood estimation under right random censoring) against the associated non-parametric Kaplan-Meier estimates (see Figure 1). We examine the two most extreme cases (i.e., smallest p-values) from the previous formal tests using the 2012/2013 data. These are the cases (Reggie Evans and George Hill) that most strongly reject the exponential distribution. We also examine the two cases (Avery Bradley and LaMarcus Aldridge) with the largest p-values. With Reggie Evans (foul level $n = 1$, $m_1 = 80$, 0 censored observations, p-value = 0.000000301) and George Hill (foul level $n = 1$, $m_1 = 76$, 9 censored observations, p-value = 0.000000681), we observe that the exponential distribution is inadequate, yet the gamma survival curves resemble the Kaplan-Meier curves. This is good because Kaplan-Meier curves are based on nonparametric methods and ought to resemble the true underlying survival distributions. Kaplan-Meier curves also readily accommodate censoring. With Avery Bradley (foul level $n = 3$, $m_3 = 37$, 11 censored observations, p-value = 0.992) and LaMarcus Aldridge (foul level $n = 2$, $m_2 = 71$, 14 censored observations, p-value = 0.995), we observe that the exponential and gamma survival curves overlap, and provide good matches to the Kaplan-Meier curves. Therefore, Figure 1 illustrates both the adequacy of the gamma distribution and its lack of drawbacks.

In the formal goodness-of-fit testing, we observed $\hat{\alpha} > 1$ in 821 out of the 1,010 datasets. This indicates an increasing hazard function for the gamma distribution. In basketball terms, this means that a player is more likely to foul when they are on the court for longer periods of time. This corresponds to our intuition that tired players are more prone to fouling. For the remainder of our investigation, we will assume fouling times obey a gamma distribution. Because the gamma distribution is a two-parameter distribution that includes the exponential as a special case, it is a more robust choice in the sense that it has more flexibility to accommodate various distributional shapes (assuming there is sufficient data to estimate the two parameters).

2.2 The Impact of Player Position and Foul Level

There is a perception that the accumulation of fouls may depend on player position. To investigate this notion, we considered NBA data from three recent seasons, 2013-2014 through 2015-2016. We further restrict the analysis to players who accumulated more than five fouls

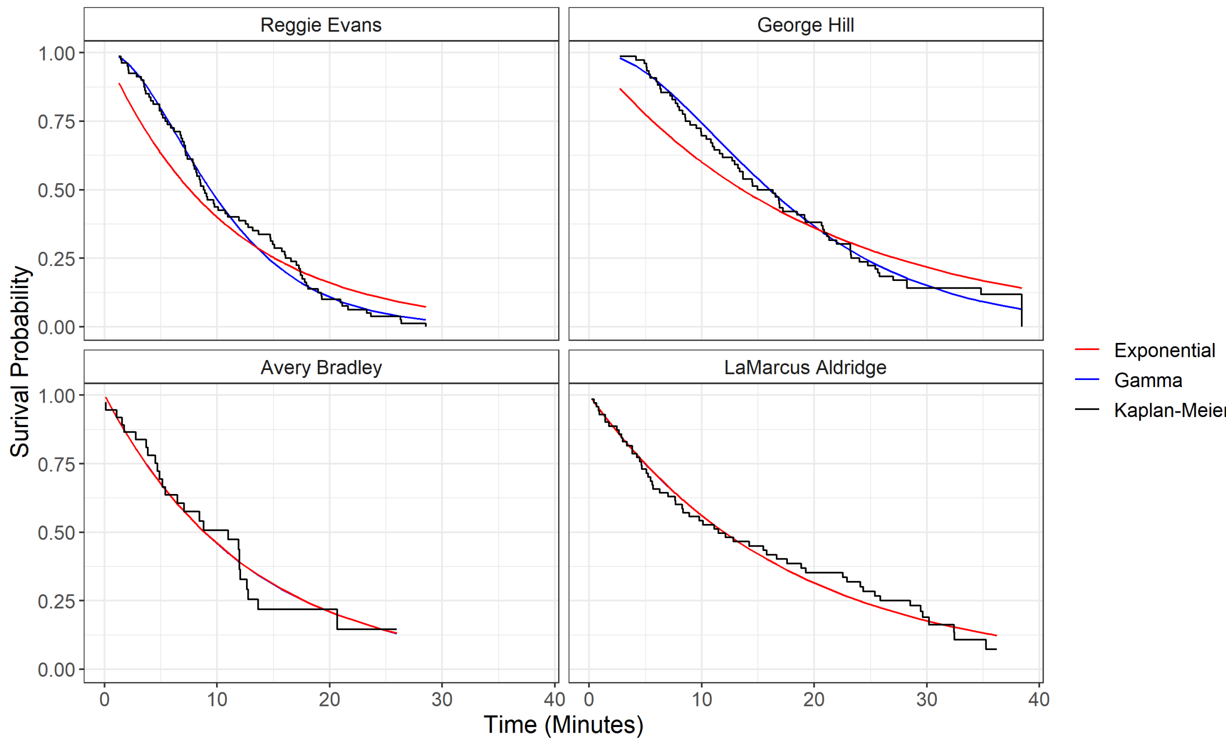

Figure 1: Estimated survival function (based on the gamma and exponential distributions) and the Kaplan-Meier estimate for four players based on data from the 2012/2013 regular season.

at each foul level and played at least 20 minutes a game for at least 60 games across the three seasons. The restriction was carried out to avoid small datasets and provide more stability in the estimates of the gamma parameters. For each foul level $n = 1, 2, \ldots, 6$, we used the maximum likelihood procedure previously discussed to estimate the gamma parameters α_{in} and β_{in} for the ith player. We classified players into one of three positions: bigs, forwards or guards according to their NBA position classification. Using www.nba.com, players are labelled as either G, G-F, F-G, F, F-C, C-F or C. We classified players according to their primary position (e.g., F-G was labelled as F). Estimated mean fouling times $\hat{\alpha}_{in}/\hat{\beta}_{in}$ were then calculated and classified according to player position. The corresponding boxplot for estimated mean fouling times according to player position is shown in Figure 2. We observe that bigs foul the most quickly, followed by forwards, and then followed by guards.

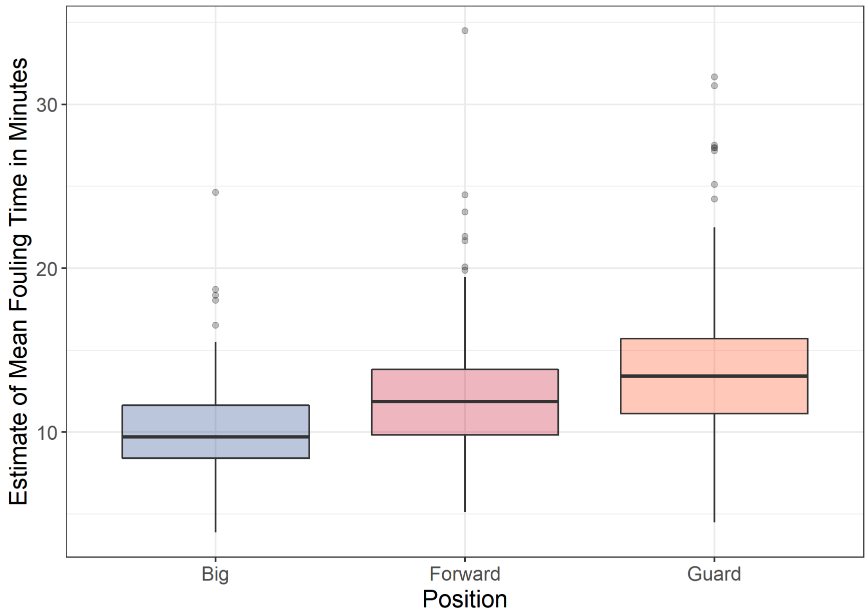

Figure 2: Boxplot of the estimated mean fouling times for each of the standard baskeball playing positions.

There is a second perception that the accumulation of fouls may depend on the foul level. We carry out a similar analysis where the estimated mean fouling times $\hat{\alpha}_{in}/\hat{\beta}_{in}$ were classified according to foul level $n = 1, 2, \ldots, 6$. The corresponding boxplot for estimated mean fouling times is shown in Figure 3. We observe a clear trend that fouls occur more quickly for increasing foul levels n.

To test the impact of player position and foul level formally, we carried out multiple two-way ANOVA analyses on the mean fouling time estimates $\hat{\alpha}_{in}/\hat{\beta}_{in}$. The ANOVA factors were playing position (bigs, forwards and guards) and foul level $n = 1, 2, \ldots, 6$. Adjusting for multiple comparisons via Tukey's HSD method, we found statistical significance for all combinations of positions; namely big-forward (p-value = 1.5e-06), big-guard (p-value = 2.2e-308) and forward-guard (p-value = 1.6e-06). We also obtained statistical significance

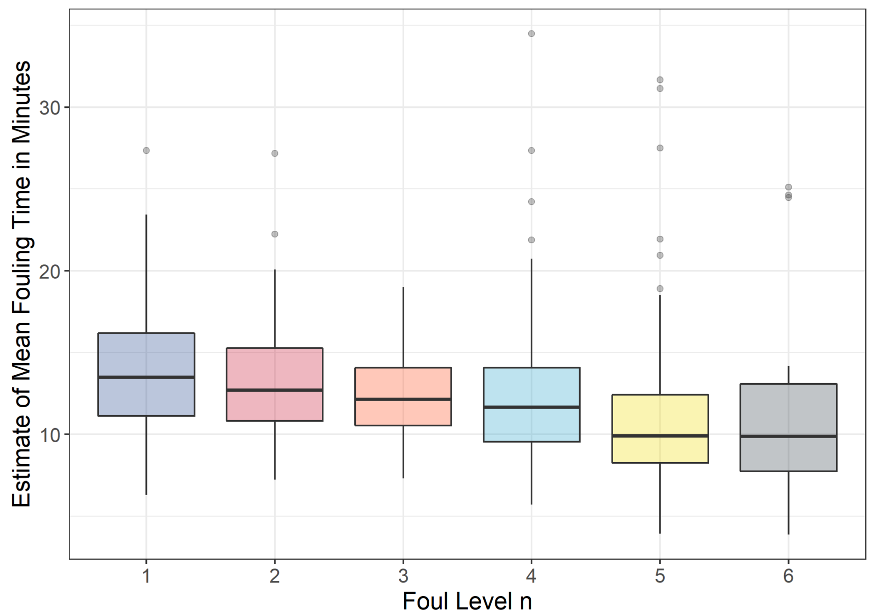

Figure 3: Boxplot of the estimated mean fouling times for foul levels $n = 1, 2, \ldots, 6$.

(p-value = 5.1e-10) when testing for the overall difference in foul levels $n = 1, 2, \ldots, 6$. However, in the 15 pairwise comparisons tests for foul level differences where we accounted for multiple comparisons, only 6 of the differences were statistically significant (1-3, 1-4, 1-5, 2-5, 3-5 and 4-5).

3 MODELLING

We use the development from Section 2.1 where α and β are vector notations for the parameters, and y and δ are vector notations corresponding to the observed data. Assuming $X_{ij}^{(n)} \sim \text{Gamma}(\alpha_{in}, \beta_{in})$, this leads to the posterior density

$$\pi(\alpha, \beta \mid y, \delta) \propto \prod_i \prod_j \prod_n f(y_{ij}^{(n)} \mid \alpha_{in}, \beta_{in})^{1-\delta_{ijn}} \, [1 - F(y_{ij}^{(n)} \mid \alpha_{in}, \beta_{in})]^{\delta_{ijn}} \, \pi(\alpha, \beta) \tag{1}$$

where f and F are the density and cumulative distribution functions corresponding to the gamma distribution and $\pi(\alpha, \beta)$ is the prior density. Here, interest concerns the unknown parameters α and β which describe the fouling time distributions.

3.1 Prior Distribution

With 75 players of interest (see Section 4) and six foul levels $n = 1, \ldots, 6$, this leads to $2(75)(6) = 900$ parameters α and β in (1). In hierarchical models, we can effectively reduce the parameterization by borrowing information between parameters. We let P_i denote the position of player i where P_i is a categorical variable with values "big" (1), "forward" (2), and "guard" (3). We have seen from the exploratory data analysis that the fouling time distributions depend on both position and foul level. We therefore consider a prior structure where $(\alpha_{in}, \beta_{in})$ arise from a distribution that depends on both the player position P_i and the foul level n.

We implemented the prior structure by imposing independence between the $(\alpha_{in}, \beta_{in})$ pairs and specifying

$$(\alpha_{in}, \beta_{in})' \sim \text{truncated\_Normal}_2((a, b)', \Sigma) \tag{2}$$

where $\Sigma = (\sigma_{ij})$. The truncations on the bivariate Normal distributions are imposed so that $\alpha_{in} > 0$ and $\beta_{in} > 0$ according to the definition of the gamma distribution. We have used some simplifying notation in (2) where it is emphasized that the hyperparameters a, b and Σ are specified and depend on the combination of the player position P_i and the foul level n.

The hyperparameters (a, b) were informed by the original 2013-2014 through 2015-2016 regular season data. At each foul level $n = 1, \ldots, 6$, we first obtained maximum likelihood estimates (MLEs) $\hat{\alpha}_{in}$ and $\hat{\beta}_{in}$ of the gamma parameters for all players who accumulated more than five fouls and played at least 20 minutes of playing time for at least 60 games across the three seasons. We then grouped the MLEs accordingly to the $3(6) = 18$ combinations corresponding to player position and foul level. The hyperparameters a and b were then determined by averaging the values of $\hat{\alpha}_{in}$ and $\hat{\beta}_{in}$ in each group. For the specification of the hyperparameter matrix Σ, we proceeded in the same fashion by calculating the second moments corresponding to $\hat{\alpha}_{in}$ and $\hat{\beta}_{in}$ in each group.

There is one exception to the hyperprior specification described above. We grouped the case $(P_i = 1, n = 5)$ with $(P_i = 1, n = 6)$, we grouped the case $(P_i = 2, n = 5)$ with $(P_i = 2, n = 6)$, and we grouped the case $(P_i = 3, n = 5)$ with $(P_i = 3, n = 6)$. This was carried out because there were fewer fouls at the higher foul levels $n = 5, 6$, and grouping provided more reliable estimation. The groupings appear justified since the mean fouling time difference between foul level 5 and foul level 6 was only 0.85 minutes (see Figure 3).

We introduced one additional feature in the prior specification based on the discovery from Section 2.3. We impose the constraint $\alpha_{i1}/\beta_{i1} \geq \alpha_{i2}/\beta_{i2} \geq \cdots \geq \alpha_{i6}/\beta_{i6}$ to reflect our knowledge that the mean fouling time decreases with increasing n. Further, if we believe that fouling time distributions have an increasing hazard function, then we may introduce the constraint $\alpha_{in} \geq 1.0$. Additional model restrictions may be useful when data used to inform the prior is not abundant.

3.2 Predictive Distributions

In the Bayesian setting, there is a convenient framework for handling predictive inference. Suppose that we are interested in the predictive distribution for the playing time $X_i^{(n)*}$ between the $(n-1)$st foul and the nth foul for player i. The density for the predictive distribution of $X_i^{(n)*}$ is given by

$$f(x) = \int f(x \mid \alpha_i, \beta_i)\ \pi(\alpha, \beta \mid y)\ d\alpha\ d\beta \tag{3}$$

where y denotes the historical data used in the determination of the posterior (1).

Fortunately, obtaining a sample from the predictive distribution (2) is a by-product of Markov chain Monte Carlo (MCMC). In the kth iteration of MCMC, we generate the parameter vector $(\alpha^{(k)}, \beta^{(k)})$. We then generate $x^{(k)} \sim f(x \mid \alpha_i^{(k)}, \beta_i^{(k)})$. Repeating the procedure, we have a sample $x^{(1)}, x^{(2)}, \ldots, x^{(N)}$ from the predictive distribution. As demonstrated in Section 4, the sample allows us to address various questions associated with fouling times.

3.3 Computation

With complex high-dimensional posterior distributions, one typically resorts to sampling-based methods to approximate posterior summaries. In this application, we use MCMC methods to generate variates from the posterior. In particular, we use the Bayesian software package Stan which is relatively simple to use and avoids the need of special purpose MCMC code. In Stan, the user only needs to specify the likelihood, the prior and the data; the determination of appropriate proposal distributions and sampling schemes are done in the background. We sample using a No-U-Turn sampler variant of Hamiltonian Monte Carlo. Stan is open source software (https://mc-stan.org) and can be accessed through RStan (https://mc-stan.org/rstan/) which is the R interface to Stan. For example, if we are able to generate variates $\alpha_{in}^{(1)}, \ldots, \alpha_{in}^{(N)}$ from the posterior (1), then $\hat{\alpha}_{in} = (1/N) \sum_{k=1}^{N} \alpha_{in}^{(k)}$ provides an estimate of the posterior mean of α_{in}.

4 RESULTS

Data were taken from the Eight Thirty Four website (Evans and Saini 2019) which consists of enhanced play-by-play data from the 2012-2013 through 2018-2019 NBA regular seasons.

Recall that the 2013-2014 through 2015-2016 NBA regular season data were used to specify the prior distribution. We now consider the posterior density (1) based on data from the 2016-2017 and 2017-2018 NBA regular seasons. We use fouling time data corresponding 75 players of interest consisting of 25 bigs, 25 forwards and 25 guards. The players selected were those who had the most minutes of playing time at their respective positions during the two seasons. From a basketball perspective, coaches care greatly about managing fouls for their star players. Players who play marginal minutes are not in danger of fouling out.

4.1 Model Validation

It is important to validate model accuracy. In particular, we compare the accuracy of our predictive distributions to predictions based on three other approaches: a baseline survival probability of 0.5, Kaplan-Meier estimation and maximum likelihood estimates for gamma distributions.

For assessing accuracy, we consider the integrated time-dependent expected Brier score as suggested by Mogensen, Ishwaran and Gerds (2012) which is adjusted for random right censoring times as described by Gerds and Schumacher (2006). In the context of survival functions, the Brier score at time t (Brier 1950) is the squared error

$$\hat{B}(t) = \frac{1}{K}\sum_{i=1}^{K}\left(1_{y_i>t} - \hat{S}(t)\right)^2$$

based on the observed failure times $y_1, \ldots, y_K$ where $\hat{S}(t)$ is the estimated survival probability and $1_{y_i>t}$ is 1 if $y_i > t$ and 0 otherwise.

Under random right censoring, the Brier score is modified using censoring weights (Gerds and Schumacher 2006). With data $(y_{ij}^{(n)}, \delta_{ij})$ and survival function $\hat{S}_i^{(n)}(t)$ for player i and foul level n, the time-dependent expected Brier Score using this adjustment is given by

$$\hat{B}_{adj}^{(n)}(t) = \frac{1}{K}\sum_{i=1}^{75}\sum_{j=1}^{m_{in}}\left(\frac{\left(0 - \hat{S}_i^{(n)}(t)\right)^2 \cdot 1_{y_{ij}^{(n)}\leq t, \delta_i = 0}}{\hat{G}(y_{ij}^{(n)})} + \frac{\left(1 - \hat{S}_i^{(n)}(t)\right)^2 \cdot 1_{y_{ij}^{(n)}>t}}{\hat{G}(t)}\right)$$

where $\hat{G}(t)$ is the Kaplan-Meier estimator of $P[C > t]$, C is the censoring time and m_{in} is the number of observations for player i at foul level n.

We calculate $\hat{B}_{adj}^{(n)}(t)$ for $t = 0, \ldots, 48$ minutes and approximate $\frac{1}{48}\int_0^{48}\hat{B}_{adj}^{(n)}(t)dt$ to obtain the integrated time-dependent expected Brier score. The integrated time-dependent expected Brier scores for each approach is calculated from out of sample predictions in the 2018-2019 NBA season from models that were fit on data from the 2016-2017 and 2017-2018 NBA seasons and are reported in Table 1.

Foul Level n	Bayesian	Kaplan-Meier	MLE	0.5 Baseline
1	0.1201	0.1200	0.1202	0.2514
2	0.1283	0.1286	0.1287	0.2553
3	0.1369	0.1371	0.1376	0.2556
4	0.1589	0.1626	0.1632	0.2372
5	0.1483	0.1563	0.1541	0.2170
6	0.1561	0.1895	0.1536	0.2273

Table 1: Integrated time-dependent expected Brier scores by foul level.

Our Bayesian model outperforms the other approaches for foul levels $n = 2, 3, 4, 5$ and is second best for foul levels $n = 1$ (Kaplan-Meier is better) and $n = 6$ (the MLE approach is better). As detailed by Redelmeier, Block and Hickam (1991) the determination of statistically significant improvements requires the true foul probabilities. Therefore, future work is needed to formally assess if these differences are statistically significant.

4.2 Extreme Foul Habits

To illustrate the variation in fouling times between players at each position, Table 2 provides the median predicted fouling times for players with extreme foul habits. We observe that players like Klay Thompson, Jimmy Butler and Marc Gasol are good at avoiding fouls whereas players like Ed Davis, Kelly Oubre Jr. and Ricky Rubio foul more quickly. This also illustrates the heavy-handedness of the $Q+1$ rule when applied to all players. It seems tactically unsound to treat Harrison Barnes or Kemba Walker in the same way as Spencer Dinwiddie or Robert Covington.

4.3 Example: Giannis Antetokounmpo - Predictive Distributions

We illustrate the fouling tendencies of Giannis Antetokounmpo of the Milwaukee Bucks based on his fouling data from the 2017-2018 and 2018-2019 regular seasons as to make the results as current as possible with the available data. We use data from the 2017-2018 and 2018-2019 regular season for the following player examples as well. Following

Player	Position	1st	2nd	3rd	4th	5th	6th
Ed Davis	C	6.0	5.7	5.5	5.1	4.9	4.4
Alex Len	C	6.3	6.0	5.5	5.7	4.7	4.0
Robin Lopez	C	12.3	10.0	7.8	7.2	5.6	4.7
Marc Gasol	C	11.3	9.6	9.0	8.4	7.5	5.4
Kelly Oubre Jr.	F	7.2	7.0	6.8	6.3	6.5	5.3
Robert Covington	F	8.2	8.0	7.1	7.3	7.0	6.1
Harrison Barnes	F	18.5	15.9	12.0	10.1	6.8	6.8
Jimmy Butler	F	20.9	15.0	11.7	9.5	6.4	6.5
Ricky Rubio	G	9.2	8.4	8.0	7.3	7.1	6.2
Spencer Dinwiddie	G	10.2	8.7	7.5	7.5	6.1	6.3
Klay Thompson	G	15.2	14.2	10.9	9.7	8.6	7.9
Kemba Walker	G	18.5	16.6	10.3	10.0	8.3	6.7

Table 2: Median predicted fouling times at each foul level for players with extreme foul habits at each position.

Section 3.2, we approximated predictive distributions for the fouling times at the foul levels $n = 1, \ldots, 6$. The estimated predictive densities are shown in Figure 4. The densities are based on 3,000 draws from the predictive distributions which are estimated by the function *geom_density_ridges* from the *ggridges* package in R. We observe that the predictive densities have long right-skewed tails indicating that there is possibility of playing a long time without fouling. This aligns with the empirical distributions at lower foul levels and differs slightly at higher foul levels where censoring is more prevalent. Like all players, we further observe that Giannis fouls quicker at later foul levels. For example, the mean predictive fouling time for Giannis is 13.2 minutes for his first foul and 11.0 minutes for his third foul.

4.4 Example: LeBron James - Endgame Scenario

LeBron James has been a star NBA player for his entire career. Any coach of LeBron would like to see him playing at the end of a match where the outcome is in the balance. Let's imagine that LeBron has picked up his fifth foul midway through the third quarter where there is 18 minutes left to play. Should LeBron's coach force LeBron to sit or should he

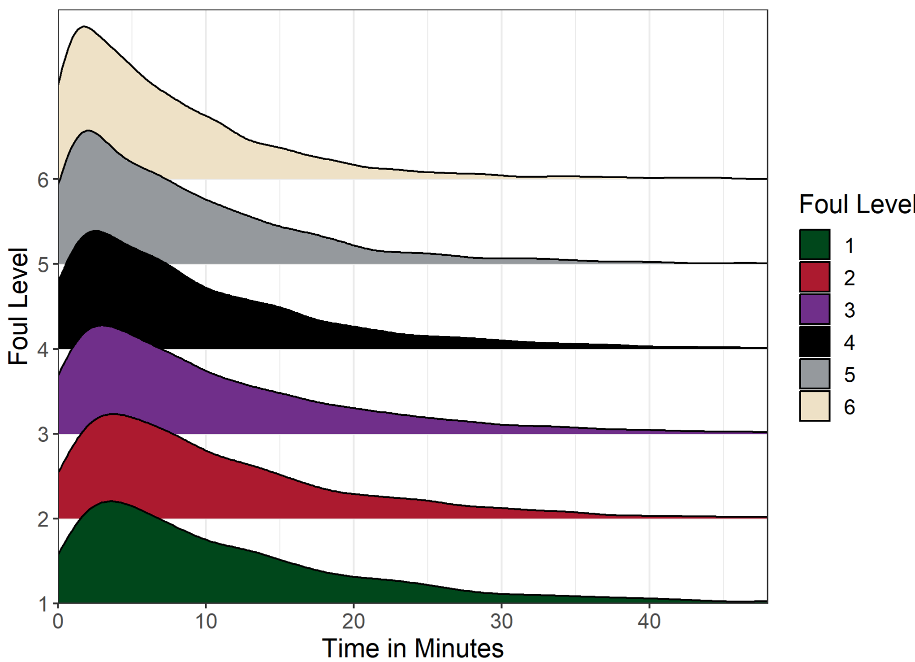

Figure 4: Predictive densities of the fouling time distribution for Giannis Antetokounmpo at the foul levels $n = 1, \ldots, 6$ based on data from the 2017-2018 and 2018-2019 regular seasons.

continue to play? Based on the MCMC output, LeBron's estimated posterior mean time for the sixth foul is 10.9 minutes (4.3 minutes longer than Giannis). However, the mean fouling time does not provide a complete picture for the problem at hand. We use the MCMC algorithm and the predictive distribution (3) to generate fouling times $x^{(1)}, x^{(2)}, \ldots, x^{(N)}$ corresponding to LeBron's sixth foul. The 10th percentile of the predictive sample based on $N = 1,000$ is 1.3 minutes. Therefore, if the coach wants LeBron playing at the end of the match with 90% probability, then the coach should force LeBron to sit and re-enter the match with 1.3 minutes remaining. This strikes us as an overly conservative strategy, where we suggest that coaches ought to be willing to have LeBron re-enter the match earlier than with 1.3 minutes remaining. For reference, the 30th and 50th percentiles for LeBron are 4.1 and 7.8 minutes, respectively. Also, note that the last time LeBron accumulated 5 fouls in a game was the 2016-2017 where this occurred twice. The first time he fouled out in less than 2.0 minutes, and the second time, in under 4 minutes.

The same scenario and analysis is considered for another impact player, Karl-Anthony Towns. Having committed his fifth foul, if his coach wants Towns playing at the end of the match with 90% probability, then the coach should force Towns to sit and re-enter the match with 1.0 minutes remaining. The 30th and 50th percentiles for Towns are 3.1 and 5.9 minutes, respectively. In the 2018-2019 season, Towns reached 5 fouls 28 times. In three of those games, the 5th foul occurred with less than 1.0 minute of playing time remaining. In four of those games, Towns fouled out in less than a minute. Therefore, while the recommendation may seem conservative, it does line up with empirical results.

We suggest that coaches understand the full distribution of outcomes and consider even larger percentiles than the 50th percentile in order to maximize minutes played. However, each individual decision must balance the threat of fouling out before the end of the game with the magnitude of minutes the player plays. Falk (2018) argues that the scale should be pushed as far as possible to maximize minutes played. We agree and want coaches to use this information to make a more informed decision.

4.5 Example: Damian Lillard - Cumulative Fouls

We provide a third player example which further illustrates the convenience of simulation-based inference using the proposed model. Following the description of the generation of predictive variates in Section 3.2, suppose we are interested in the total time T that Damian Lillard can play following his second foul. If x_j is the predicted time between the $(j-1)$st foul and the jth foul, then our interest concerns $T = x_3 + x_4 + x_5 + x_6$.

In Figure 5, we provide the survival curve corresponding to T for Damian Lillard. We observe that Lillard's median time for fouling out exceeds 48 minutes (ie. the length of a match). Therefore, in the case of Damian Lillard, it may be unnecessarily cautious for coaches to follow the $Q + 1$ rule.

If Damian Lillard picks up his second foul 8 minutes into the first quarter, Lillard would be considered to be in foul trouble according to conventional wisdom. However, we would estimate that there is only a 64% chance of surviving if Lillard were to play the final 40 minutes of the game. Of course, this is unrealistic since Lillard needs some rest throughout the game. However, Lillard would have an 88% chance of playing his season average of 36 minutes in the game (an additional 28 minutes given the 8 minutes that he has already played). Therefore, we would not consider Lillard to be at great risk of fouling out and we would not recommend changing his regular substitution rotations.

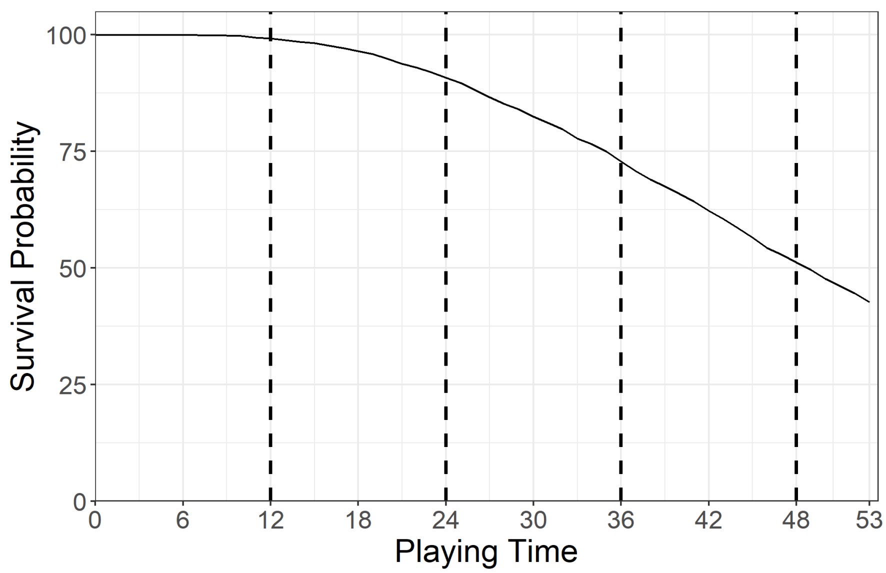

Figure 5: Survival curve of the total playing time $T = x_3 + x_4 + x_5 + x_6$ prior to fouling out for Damian Lillard after he has committed his second foul.

5 DISCUSSION

This paper introduces parametric models in a Bayesian framework for the analysis of fouling time distributions. The problem is important since NBA players foul differently, and coaches wish to manage playing minutes well.

Some of the messages in this paper include (1) that the gamma distribution provides a flexible and appropriate distribution for fouling times, (2) that mean fouling times decrease across the positional types given by bigs, forwards and guards, and (3) that mean fouling times decrease as more fouls are accumulated. Future work may consider the impact of consecutive playing time versus segmented playing time, and other covariates such as those suggested by Evans (2017a, 2017b).

It is our hope that the methods presented here may help NBA teams make better substitution decisions. Should teams implement the methods, we suggest that they remove intentional fouls from the dataset. Although intentional fouls are infrequent, they should not be included as they do not characterize individual fouling behaviour.

6 REFERENCES

Beaudoin, D. and Swartz, T.B. (2010). "Strategies for pulling the goalie in hockey," *The American Statistican*, 64, 197-204.

Brier, G.W. (1950). "Verification of forecasts expressed in terms of probability," *Monthly Weather Review*, 78, 1-3.

D'Agostino, R.B. and Stephens, M.A. (1986). *Goodness-of-Fit Techniques*, Marcel Dekker, New York.

Davis, N. and Lopez, M. (2015). "NHL coaches are pulling goalies earlier than ever," *FiveThirtyEight*, accessed September 20, 2019 at https://fivethirtyeight.com/features/nhl-coaches-are-pulling-goalies-earlier-than-ever/

Evans, K. and Saini U. (2019). "Eight Thirty Four," accessed November 11, 2019 at https://eightthirtyfour.com/data

Evans, K. (2017a). "NBA Fouls - I love DeMarcus Cousins," accessed September 20, 2019 at https://causalkathy.com/2017/01/31/i-love-demarcus-cousins/

Evans, K. (2017b). "NBA Fouls - Survival Analysis," accessed November 3, 2019 at https://causalkathy.com/2017/02/21/nba-fouls-survival-analysis/

Falk, B. (2018). "The Trouble with Foul Trouble," *Cleaning The Glass*, accessed September 22, 2019 at https://cleaningtheglass.com/the-trouble-with-foul-trouble/

Gerds, T. and Schumacher, M. (2006). "Consistent estimation of the expected Brier score in general survival models with right-censored event times," *Biometrical Journal*, 48(6), 1029–1040.

Harter, H.L. and Moore, A.H. (1965). "Maximum-likelihood estimation of Gamma and Weibull populations from complete and censored samples," *Technometrics*, 7, 639-643.

Kalbfleisch, J.D. and Prentice, R.L. (2002). *The Statistical Analysis of Failure Time Data, Second Edition*, John Wiley & Sons, Inc., Hoboken, New Jersey.

Klobuchar, A. (2018). "Basketball foul analysis: A look at the inefficiency in current NBA substitution-after-fouls behavior," *Aidan's Musings*, accessed September 22, 2019 at http://aidanklobuchar.com/home/data-based-writings/fouls/

Lewis, M. (2013). *Moneyball: The Art of Winning an Unfair Game*, WW Norton, New York.

Maymin, P., Maymin, A. and Shen, E. (2012). "How much trouble is early foul trouble? Strategically idling resources in the NBA," *International Journa of Sport Finance*, 7, 324-339.

Mogensen, U.B., Ishwaran, H. and Gerds, T.A. (2012). "Evaluating random forests for survival analysis using prediction error curves," *Journal of Statistical Software*, 50(11), 1–23.

Rochford, A. (2017). "NBA foul calls and Bayesian item response theory," accessed September 22, 2019 at https://austinrochford.com/posts/2017-04-04-nba-irt.html

Redelmeier, D.A., Bloch, D.A. and Hickam D.H. (1991). "Assessing predictive accuracy: how to compare Brier scores," *Journal of Clinical Epidemiology*. 44(11), 1141-1146.

Yam, D. and Lopez, M. (2018). "Quantifying the causal effects of conservative fourth down decision making in the National Football League," Available at http://dx.doi.org/10.2139/ssrn.3114242