Project Proposal

Tom Kennon

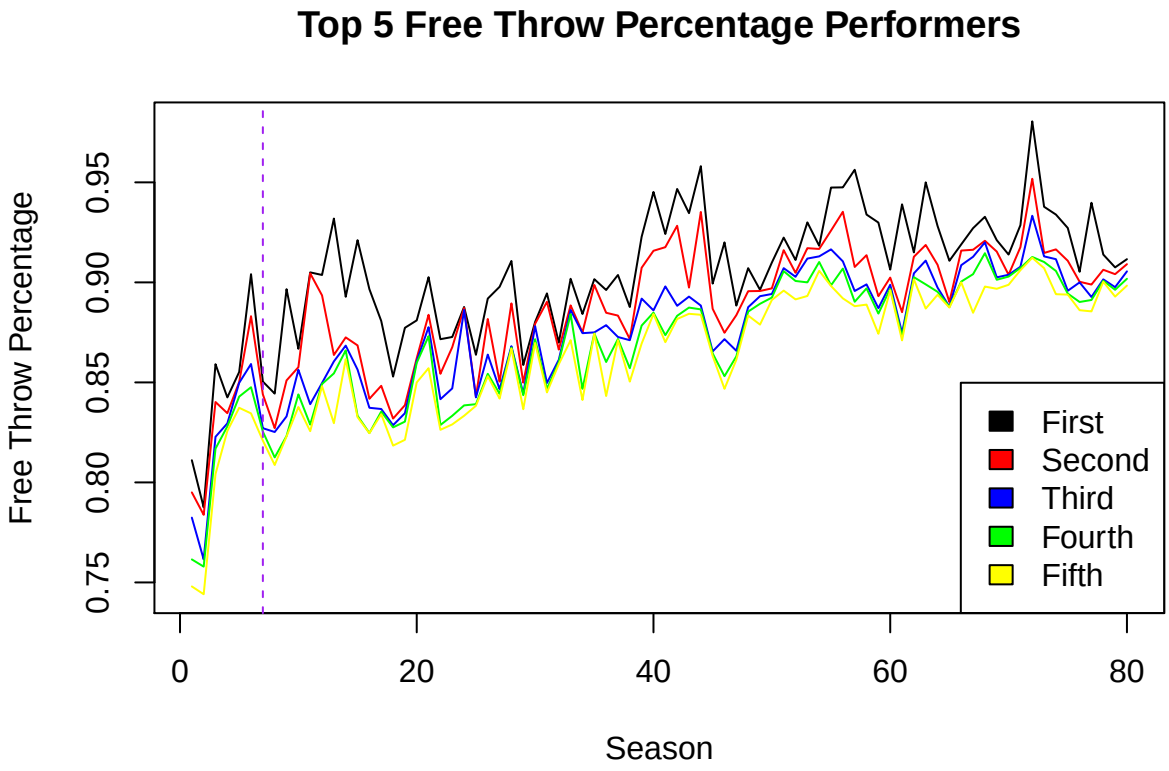
February 20, 2018

The goal of this project is to determine if there is a limit for an NBA player's free throw percentage in a season. Free throws are one of the most unique opportunities in sports where a player is given a "free" uncontested attempt to score. Ideally, a player should make every free chance they get by shooting 100% of their free throws, yet no NBA player has ever completed this feat for a whole season. With NBA players shooting better and better each season as of recently, a burning question for many NBA fans is: Is it possible to achieve a perfect free throw season making every shot one takes? I will use extreme value theory estimation to determine the best conceivable free throw percentage for an NBA player in a season.

The dataset was taken from Basketball Reference (https://www.basketball-reference.com/). This website is a comprehensive source for various statistics in basketball's history, most specifically the NBA. I pulled the top 10 NBA individual players' free throw percentages for each season from 1947-2017 (80 seasons total). Here is a preview of the dataset:

```
##    Season  Lg         X1st  perc1         X2nd  perc2         X3rd  perc3
## 1 2016-17 NBA C. McCollum  0.9116   I. Thomas  0.9091   K. Irving  0.9055
## 2 2015-16 NBA    S. Curry  0.9075 J. Crawford  0.9041   K. Durant  0.8976
## 3 2014-15 NBA    S. Curry  0.9139    J. Meeks  0.9063   J. Redick  0.9015
## 4 2013-14 NBA  B. Roberts  0.9398 D. Nowitzki  0.8989  R. Jackson  0.8927
## 5 2012-13 NBA   K. Durant  0.9053    S. Curry  0.9003   J. Redick  0.9000
## 6 2011-12 NBA J. Crawford  0.9272   J. Redick  0.9108 D. Nowitzki  0.8958
```

I have begun exploratory analysis on this dataset. Here is a plot of the top 5 free throw percentages from each of the 80 NBA seasons in this dataset.

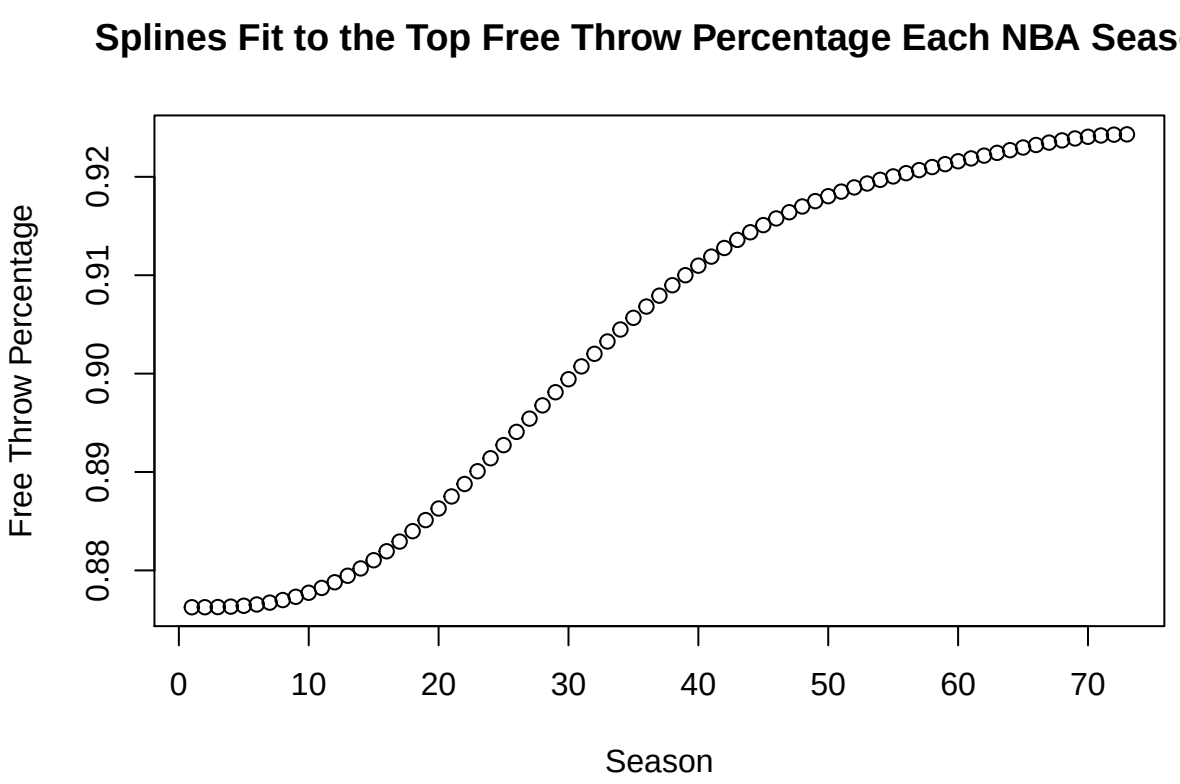


When the NBA first began in the late 1940s to early 1950s, players were still adjusting to the league's rules so there is a rapid improvement in NBA free throw percentages that is unlikely to happen again. However, there is still a slight,fairly stable improvement in free throw percentages from then onward. My goal is to model this upward trend and determine if there is a top limit besides the obvious 1.00. I will attempt to use the generalized extreme value (gev) distribution.

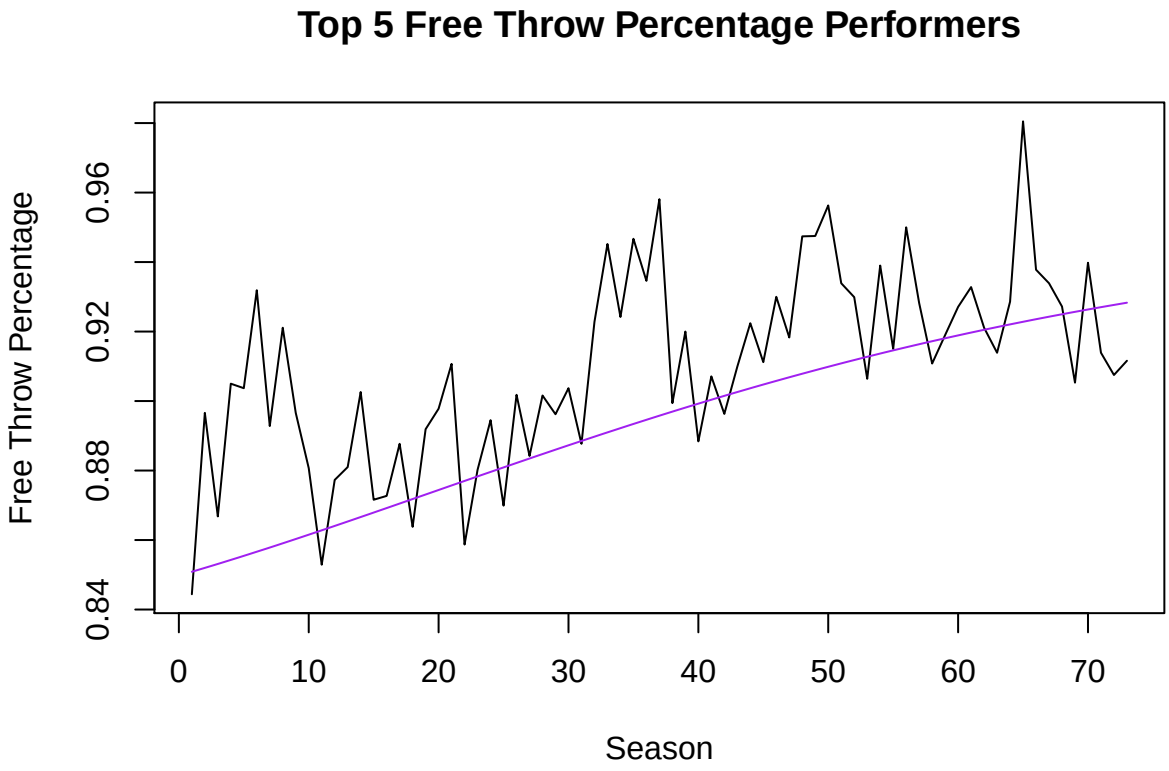
The pdf is defined as: $f(x) = \frac{1}{\sigma}\left(1+\xi\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}-1} e^{-\left(1+\xi\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}}}$.

The generalized extreme value (gev) distribution takes three parameters: location $\mu(t)$, scale σ, and shape ξ. The location parameter depends on time (how many seasons we use). This is the key parameter to model. I plan on using and evaluating the fit of various non-linear implementations for this location parameter including a splines approximation and a gompertz curve.

The splines approximation approach uses the {splines2} package written by Wenjie Wang and Jun Yan. This technique takess in a degrees of freedom value and number of knots to approximate a polynomial curve. I can fit a splines matrix as the location parameter $\mu(t)$ of the gev distribution to fit to the free throw data. Below is the splines fitting to the top 1 NBA free throw percentage of each season.



The gompertz approach uses the gompertz function defined as: $f(t) = ae^{-be^{(-ct)}} + z$ which is a function of time t. a is an asymptote, b describes where the curve is placed on the x axis , c is the growth rate, and z is the intercept. I can fit a gompertz curve as the location parameter $\mu(t)$ of the gev distribution to fit to the free throw data. Below is an example of a gompertz function implemented as the location parameter in the gev distribution fitted to the NBA free throw data.



I will then evaluate the estimated limit on NBA free throw percentage in an NBA season including confidence intervals. I will compare the different approaches and evaluate the fits. I will then interpret what these results mean for NBA fans.