

NBA Rookie Career Projections

Intro

- The data contains 20 different statistics on over 1,300 different NBA rookies
- 19 Variables, including points per game, number of games played, field goal percentage, target reached, seals, assists, free throw percentage, etc.

Goal:

- Determine the predictive probability a rookie will reach the target of playing for at least 5 years in the NBA, based on the variables **Field Goal Percentage** and **Games Played**
 - Field Goal Percentage** - the percentage of shots made during the game out of the total number of shots they have taken
 - Games Played** - number of games played, (starts and coming in off the bench)
 - Target** variable is binomial: 1 for yes, played for at least 5 years; 0 for no, did NOT play for at least 5 years

Methods:

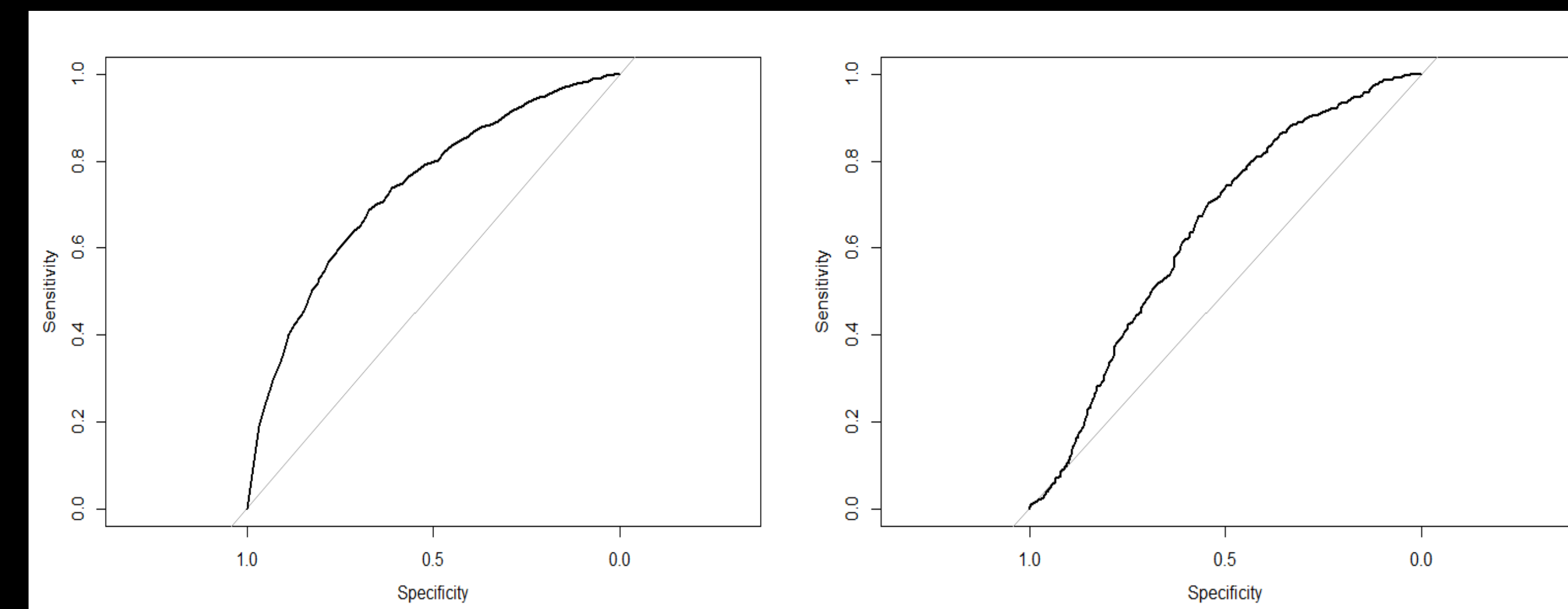
- Generalized Linear Model:

$$E[Y] = g^{-1}(X\beta).$$

- Systematic Component : $(X\beta)$, this specifies the explanatory variables in the model
- Random Component : the probability of the response variable, this can either be a normal distribution for Y with linear regression or a binomial distribution for Y in the binary logistic regression.
- Link Function : denoted by g^{-1} . This is how the expected value of the response relates to the linear predictor variables.

Methods(continued)

Receiver Operative Characteristic Curve



This **tests if the model is a good fit**, null hypothesis is the fit of the model is good, vs. the alternative hypothesis that the fit of the model is NOT good. Grouping the observed and predicted values, then comparing the observed and predicted using a goodness of fit test.

- This **model is a good fit**, area under curve is between 0.5 and 0.85. Left curve, .73359 (**Games Played**), Right curve, .6422 (**Field Goal Percentage**)
- Logistic Regression due to binomial variable

$$\ln\left(\frac{\pi}{(1-\pi)}\right)$$

Odds : the ratio of the probability of success to the probability of failure, denoted in formula as

$$\frac{\pi}{(1-\pi)}$$

Analysis/Conclusion:

- Odds are hard to interpret/understand.
- Odds were converted to Predictive probability using $\frac{\psi}{1+\psi}$
 ψ represents the odds
- Predictive probability of a player reaching the target goal of playing 5 or more years in the NBA. Using Logistic Regression

	23.8	40.2	44.1	47.9	73.7
11	0.482	0.521	0.531	0.540	0.6
47	0.572	0.610	0.619	0.627	0.683
63	0.611	0.648	0.656	0.664	0.716
77	0.644	0.679	0.887	0.695	0.744
82	0.656	0.690	0.698	0.705	0.754

- As **field goal percentage of their rookie increases, so does the probability that a player will play for at least 5 years** with the number of games played held constant.
- As **the number of games played in their rookie year goes up, so does the probability that the player will play for 5 years** with field goal percentage held constant.

A rookie has a higher probability of playing for at least 5 years in the NBA, if they play in more games than the other rookies and/or have a higher field goal percentage than other rookies.

References:

"6.1 - Introduction to Generalized Linear Models." 6.1 - Introduction to Generalized Linear Models | STAT 504, online.stat.psu.edu/stat504/node/216/.