Predicting Fantasy Football with 2-Step Bayesian Hierarchical Model

By: Michael Sarkis

Executive Summary

The popular game of fantasy football is a great candidate for Bayesian hierarchical modeling. The following analysis used 8 years of data and 3000 total players to predict future yearlong fantasy scores based on last year performance. Using cross validation and an iterative modeling process a two-part model was developed that factored in several variables including position, team and salary among others. First a logistic model was used to evaluate whether a player was likely to score "zero" points next year and then all other players' scores were predicted with a Bayesian hierarchical model. The models fit was average with projections having a root mean squared error of 62.6 (for values ranging from 0 to 350). This model provides another useful tool to be used in evaluating fantasy football players.

Introduction

Fantasy football is a popular game with over 46 million players and $22 billion dollars in profit in the US in 2020.[1] It allows fans of football to create a fantasy team of football players and compete with this team against others. There are two main types of fantasy football, daily fantasy (mainly Draft King or Fan Duel) and traditional yearlong fantasy (mainly ESPN or NFL). This analysis focuses on the traditional yearlong fantasy as it is the more popular version of the game. The basic concept of fantasy football is that a fantasy player wants their team to score more fantasy points than their opponent's team in a given week during the season. Points are awarded for many different things, but are meant to represent an NFL player's contribution to their game. For example, a player gets 6 fantasy points for a receiving or rushing touchdown and 1 fantasy points for every 10 yards of rushing or receiving. It should be noted that for this analysis ESPN's standard point per reception scoring is used. The overall goal of this analysis is to use a Bayesian hierarchical model to accurately predict a player's fantasy points production for a season given their past season's results.

Data

The player game data used in this analysis was scraped from Pro Football Reference, a publicly available source of past NFL statistics and fantasy scores. All fantasy players from 2012 to 2020 were used for a total of just over 4000. For salary information was scraped from overthecap.com a publicly available source of historical NFL contracts. After combining salary data with game data over 1000 players had to be excluded due to missing information. However, 3000 players remained with over 250 for each season, which is more than the usual 200 players which are generally relevant for any fantasy league. The data was also restricted to the positions of

---

[1] Fantasy Sports & Gaming Association - https://thefsga.org/fsga-faqs/

quarterback (QB), running back (RB), wide receiver (WR) and tight end (TE). The fantasy points from last season and the salary data were all centered and scaled.

Methods

This analysis was done using a hierarchical logistic regression followed by a Bayesian hierarchical model. The logistic model process began by creating several candidate models that were then tested with 5-fold cross validation to see which of the models had the best prediction accuracy. Then a similar several Bayesian hierarchical models were tested against one another using root mean squared error (RMSE) compared to the true fantasy score for the next year. The models varied by a single fixed effect or grouping to evaluate if specific variables were important for improving the error and accuracy.

Results

The final models ended up including the grouping variables position and the fixed effects of age, average salary per year (APY) and previous fantasy points, while the second model also included grouping by position. The model specifications are listed below:

$$IsZero_{ip}|\mathrm{x}_{ip} \sim Bernoulli(\pi_{ip}),\ i = 1, \ldots, n;\ \ p = 1, .., 4$$

$$log\left(\frac{\pi_{ip}}{1-\pi_{ip}}\right) = \beta_0 + b_{1p} + \beta_1 PFP + \beta_2 APY + \sum_a \beta_{3a} I(Age = a)$$

$\beta_0 = \text{Intercept}$

$b_{1p} = \text{Random Effect of Position};\ b_{1p} \sim N(0, \sigma^2)$

$\beta_1 = \text{Fixed Effect for Previous Year's Fantasy Points}$

$\beta_2 = \text{Fixed Effect for Average Salary Per Year}$

$\beta_{3a} = \text{Fixed Effect for Age Categories}$

$$(Y_{ipt}|IsZero_{ip} = 0) = \mu + b_{1p} + b_{2t} + \beta_1 PFP + \beta_2 APY + \sum_a \beta_{3a} I(Age = a) + \epsilon_{ipt}$$

$$(Y_{ipt}|IsZero_{ip} = 1) = 15$$

$\mu = \text{Overall Mean};$ $\qquad Prior: \mu \sim N(100, 40)$

$b_{1p} = \text{Random Effect of Position};\ b_{1p} \sim N(0, \tau_p^2);$ $\qquad Prior: \tau_p^2 \sim IG(2, 10)$

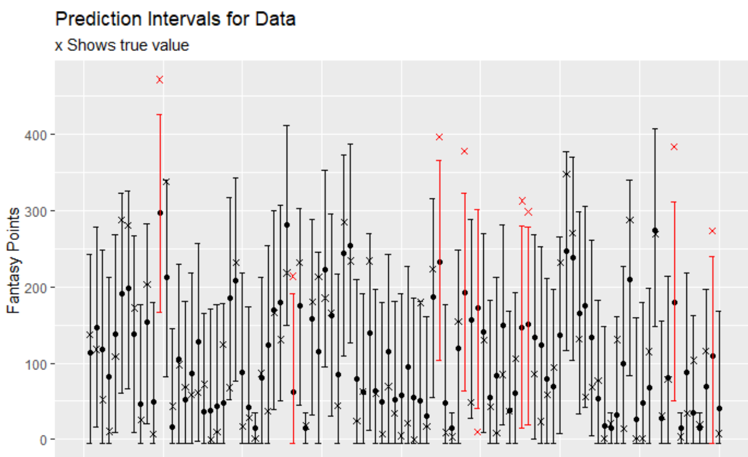
$b_{2t} = \text{Random Effect of Team};\ b_{2t} \sim N(0, \tau_t^2)$ $\qquad Prior: \tau_t^2 \sim IG(2, 5)$

$\epsilon_{ipt} \sim N(0, \sigma^2) \perp b_{1p}, b_{2t}$ $\qquad Prior: \sigma^2 \sim IG(2.5, 150)$

The first model had an overall accuracy of 82%, but more importantly, only incorrectly predicted an observation would be zero 2.03% of the time. However, it was ultimately found that including this logistic regression had little effect on the final RMSE as a modeling approach using only the second model had an RMSE that was only .02 worse.

For the second model the prior for the mean was chosen to be somewhat centered around average fantasy values, while the other priors were chosen to be uninformative. The position random effects are shown in the appendix along with the team random effects. For the fixed effects the coefficient for each value of age decreased as the age got higher, except for the highest age grouping. The coefficient for the previous year's fantasy points was .62, which shows players' year to year performance is somewhat consistent. Lastly the effect of the APY variable was .30 for every $100,000.

A random sample of the 100 of the predictions from this model are shown below. Many of the prediction intervals are very large, but most of them do encompass the true value. Also, all



prediction intervals lower bounds were limited to -5 as no fantasy player ever scores below that value. As a naïve comparison for this model, I found the RMSE of using last season's fantasy score as the prediction for next season's fantasy score. The RMSE of the naïve prediction was 72.71, while this model has an RMSE of 62.60. This represents about a 14% reduction in RMSE from the naïve approach. It should also be noted that fantasy scores generally range from 0 to 350 to understand the scale of these RMSEs.

Conclusion

Ultimately this analysis is a good basic model for evaluating noisy fantasy football data. The hierarchical model is a useful approach that helps to quantify the team and positional effects. The logistic stage of the model was lacking and did not appear to improve the RMSE by much if at all. Also, the final model's predictions still inaccurate, but there was a significant improvement over the naïve approach. For future analysis a different approach to the large number of zeros and low values in the response data could be used. Additionally, using game specific data (snap counts, etc.) or schedule strength data may be helpful. Future modeling could also include a yearly fantasy drafting model based on the yearly projects or an expansion to weekly fantasy projections.

Appendix

Appendix Fig. 1

The logistic regression cutoff point included the majority of the observations clustered near 0.

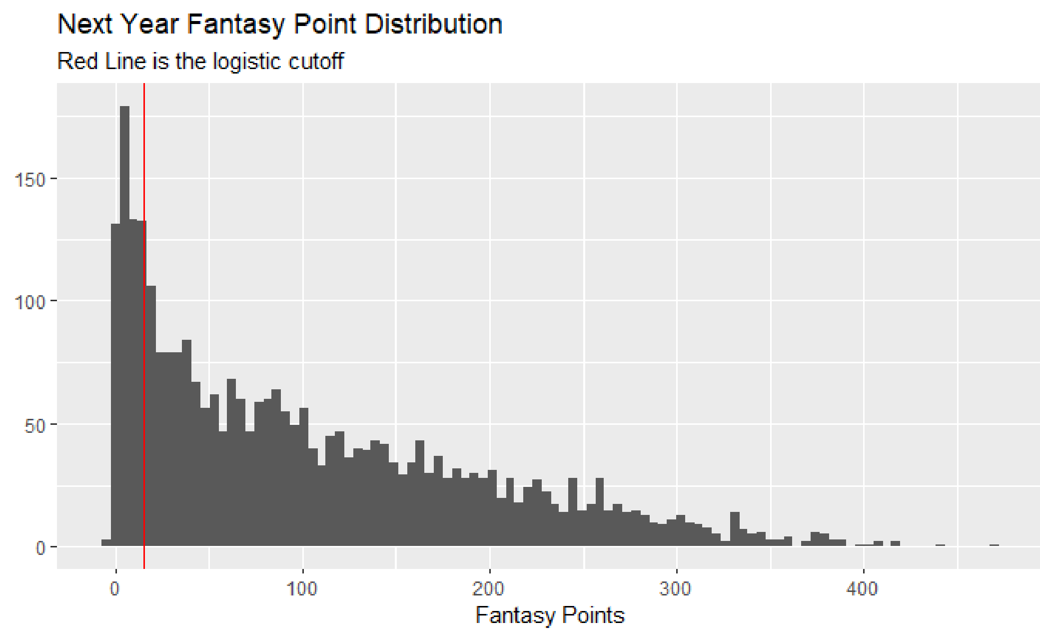
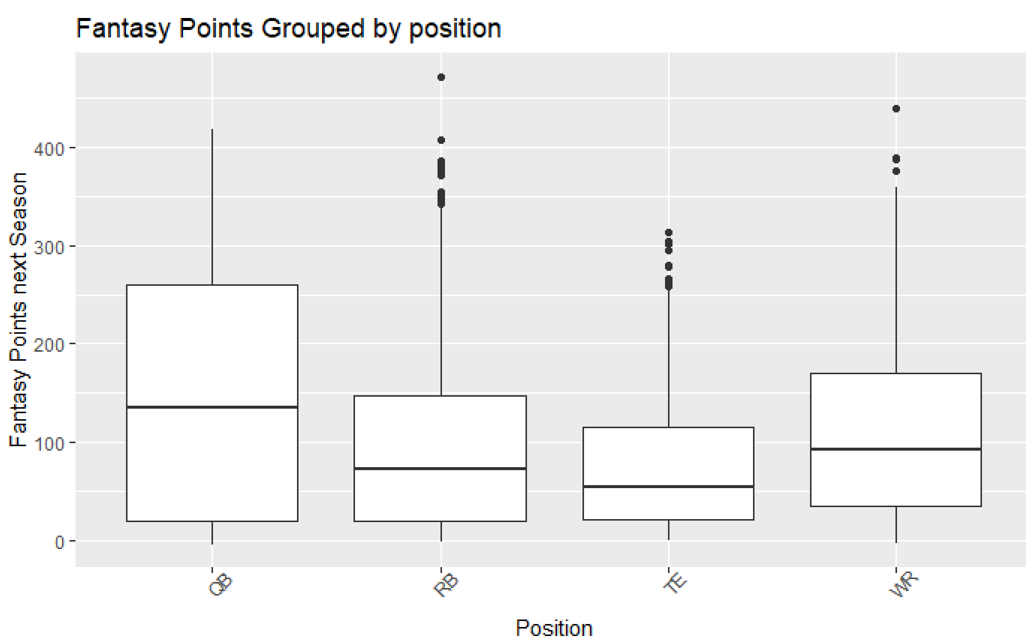


Fig 1

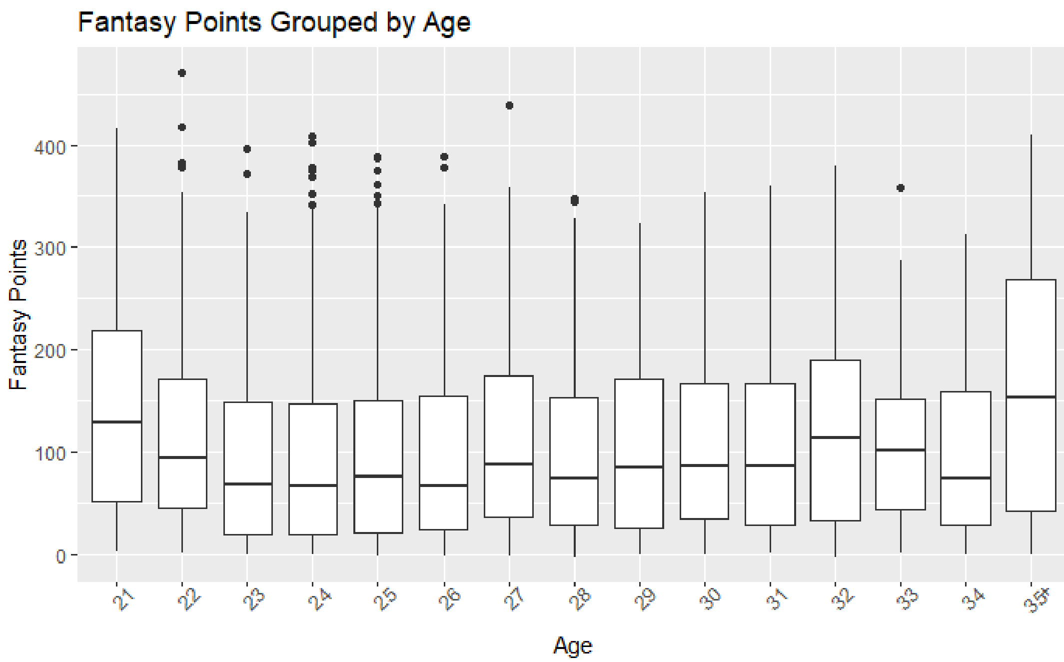
Appendix Fig. 2

Heterogeneity in fantasy points by position. This follows standard fantasy knowledge that quarterbacks score the most fantasy points.



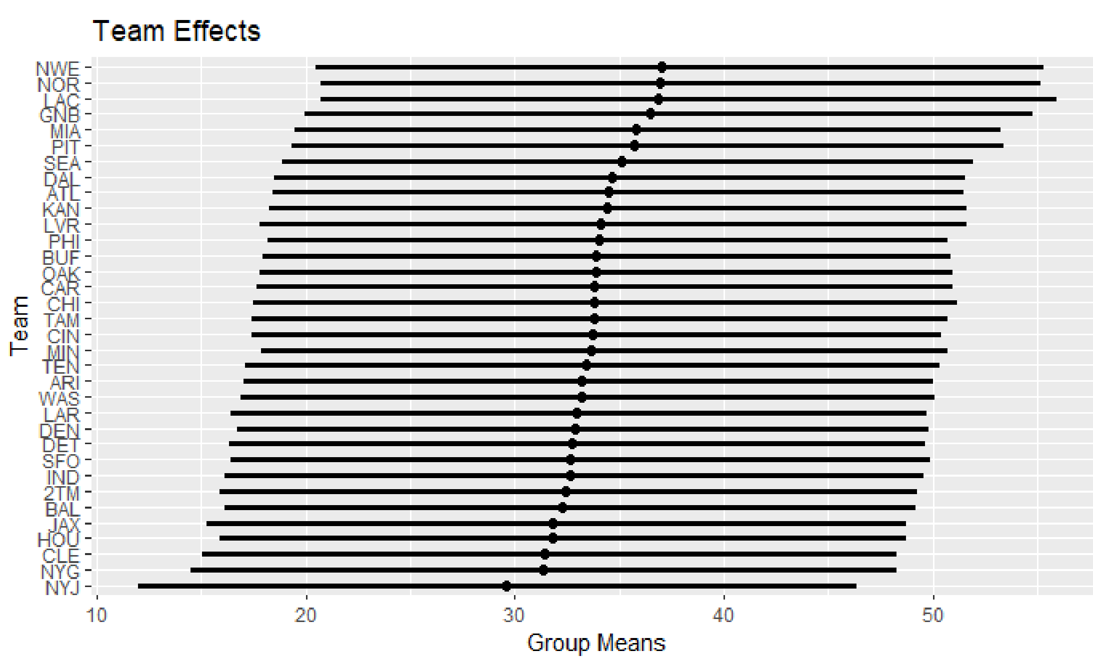
Appendix Fig. 3

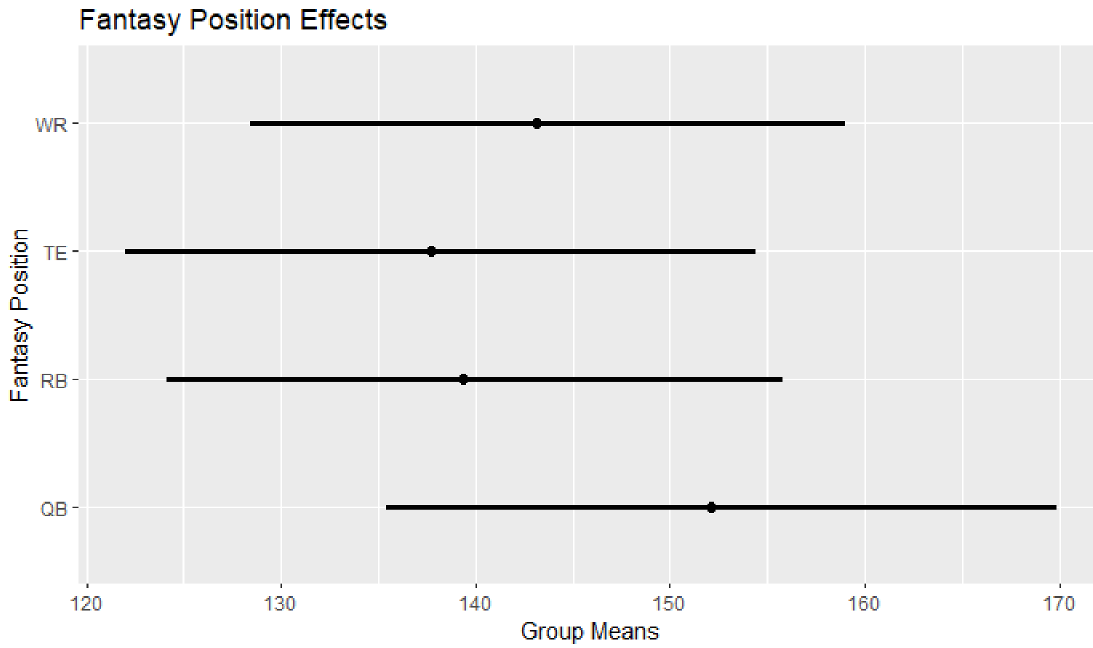
Age is treated as a categorical variable for the analysis as the youngest and oldest players have increased fantasy scores.



Appendix Fig. 4 & 5

The team effects shown below follow closely with the exploratory data analysis done. As do position effects, except TE which is higher than expected.




Fantasy Position Effects
WR
TE
RB
QB
Fantasy Position
120
130
140
150
160
170
Group Means