

Bayesian Logistic Regression

Matej Miočić, 63180206, mm9520@student.uni-lj.si

I. INTRODUCTION

In this homework we were tasked with using Bayesian Logistic regression to infer the relationship between the input variables and shot success. We look at the beta coefficients and we show results and their meaning.

II. RESULTS

Before running the results we think about what angle and distance mean for scoring hoops. If we think intuitively, we predict that the larger the distance the less chance we score, that means beta coefficients for distance should be definitely negative (the larger the distance, more samples with Made = 0). We are doing MCMC sampling, so the samples should be distributed normally. So the mean would be a negative number and standard deviation would depend on the number of samples we take.

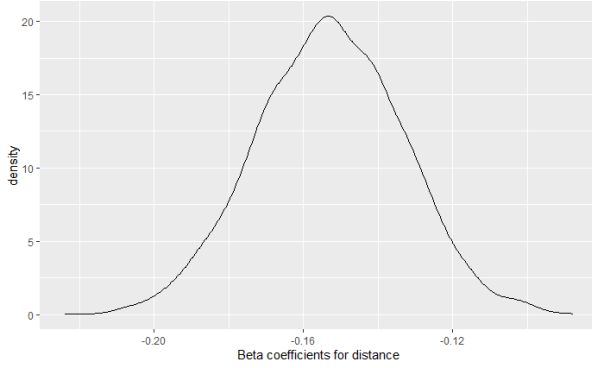


Figure 1. Density plot of beta coefficients for distance.

After observing the results in Figure 1, we see that our predictions were correct since beta coefficients density resembles a normal distribution with negative mean.

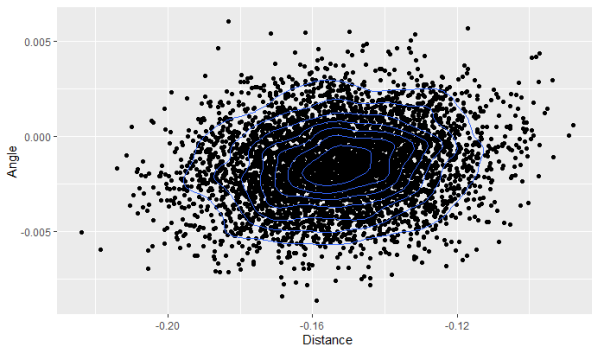


Figure 2. Scatter plot of beta coefficients for the whole dataset.

First thing we observe in Figures 2 and 3 is that some of the beta coefficients for distance are positive when looking at the small subset. Since we are sampling with only 50 instances, some samples contain coefficients for distance that tell us that more shots are successful when distance is increasing which is intuitively wrong for real world, but for this smaller dataset it might be true. Meanwhile beta coefficients for angle are more

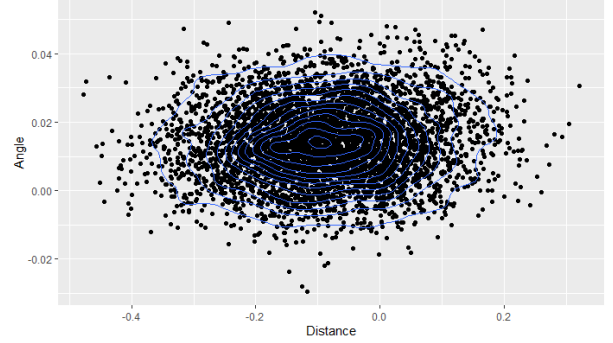


Figure 3. Scatter plot of beta coefficients for the subset of 50 instances.

similar between the 2 datasets. Mean of angle coefficients is positive in the smaller subset (0.014), where as mean for the whole dataset is negative (-0.002), but still the means are very close together.

Since every absolute value of beta coefficient for distance is larger than the absolute value of beta coefficient for angle, we conclude that distance is more important for predicting whether a shot was a success.

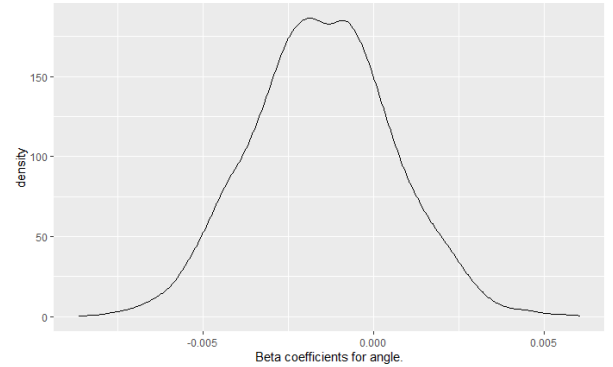


Figure 4. Density plot of beta coefficients for angle.

In Figure 4 we observe that the majority of coefficients are negative, we would say that shot success decreases with increasing angle, but we should also point out that these coefficients are extremely small, so it means that it does not impact the shot success as much relative to the distance.

III. CONCLUSION

We explored dataset of 1000+ basketball shots with Bayesian Logistic Regression. We have shown results of beta coefficient and have proven and disproven some initial thoughts about distance and angle impact on shot success.