
# 2008 NFL REGULAR SEASON FIELD GOAL ATTEMPTS

## SIMPLE LOGISTIC REGRESSION (WEEK 6)

**BY:ANDREW HENEGHAN**

# INTRODUCTION

- Simple logistic regression
- Success probability curves
- Prediction of the ability of making a field goal attempt based off kicking distance.

# SIMPLE LOGISTIC REGRESSION MODEL

- The structure of model is like that of a simple linear regression.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \epsilon,$$

- The response variable, Y, is binary.
  - It has only two values, such as “success” or “failure.”

- The probability of “success” outcome (success probability) is p.
  - p is also described as the expected value of the binary random variable.

$$p = E[Y] = \beta_0 + \beta_1 x$$

- The X variable acts as a factor that can greatly impact the success probability.

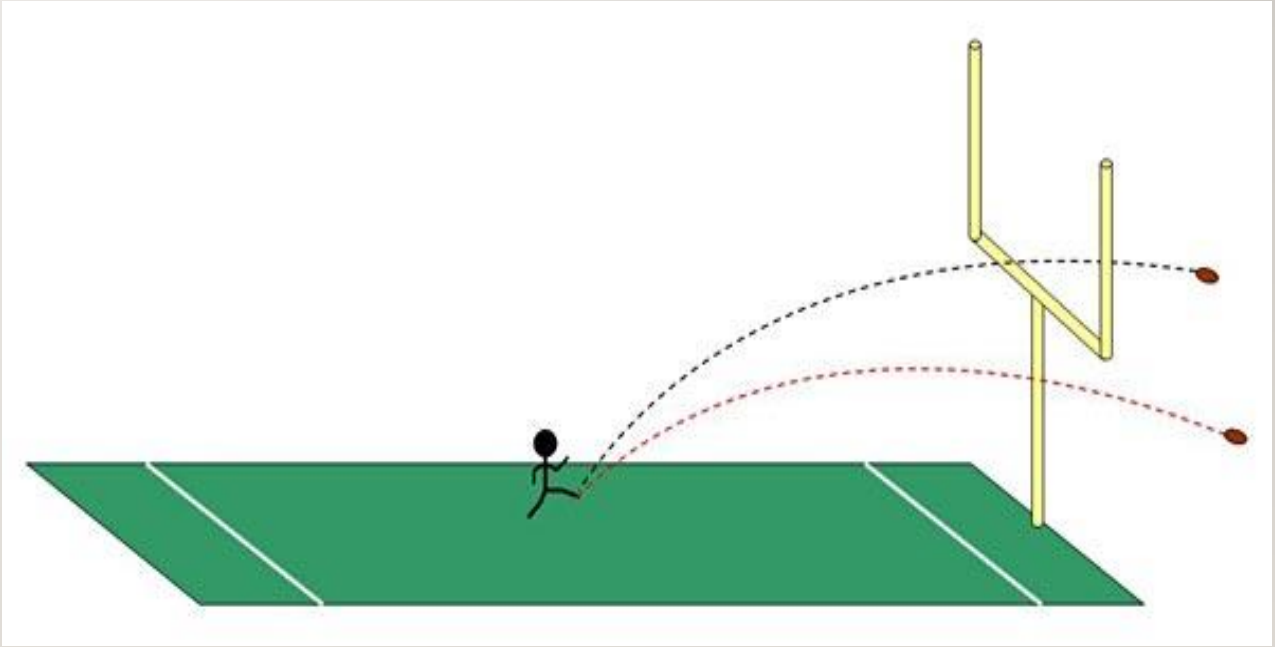
# THE “NFL FIELD GOAL ATTEMPTS BY KICKER WITH SITUATIONAL INFO” DATA SET

- Data was collected from previous results of NFL regular season field goal attempts.
- Collected during the 2008 season.
- The Elias Sports Bureau (ESB) wanted to look at different kinds of factors affecting likelihood of a successful field goal attempt.
- Data set contains 1039 observations on 23 different variables.

# THE VARIABLES

**GOOD**: the binary response variable, whether the field goal attempt is made or not. (1 if Success, 0 if Miss)

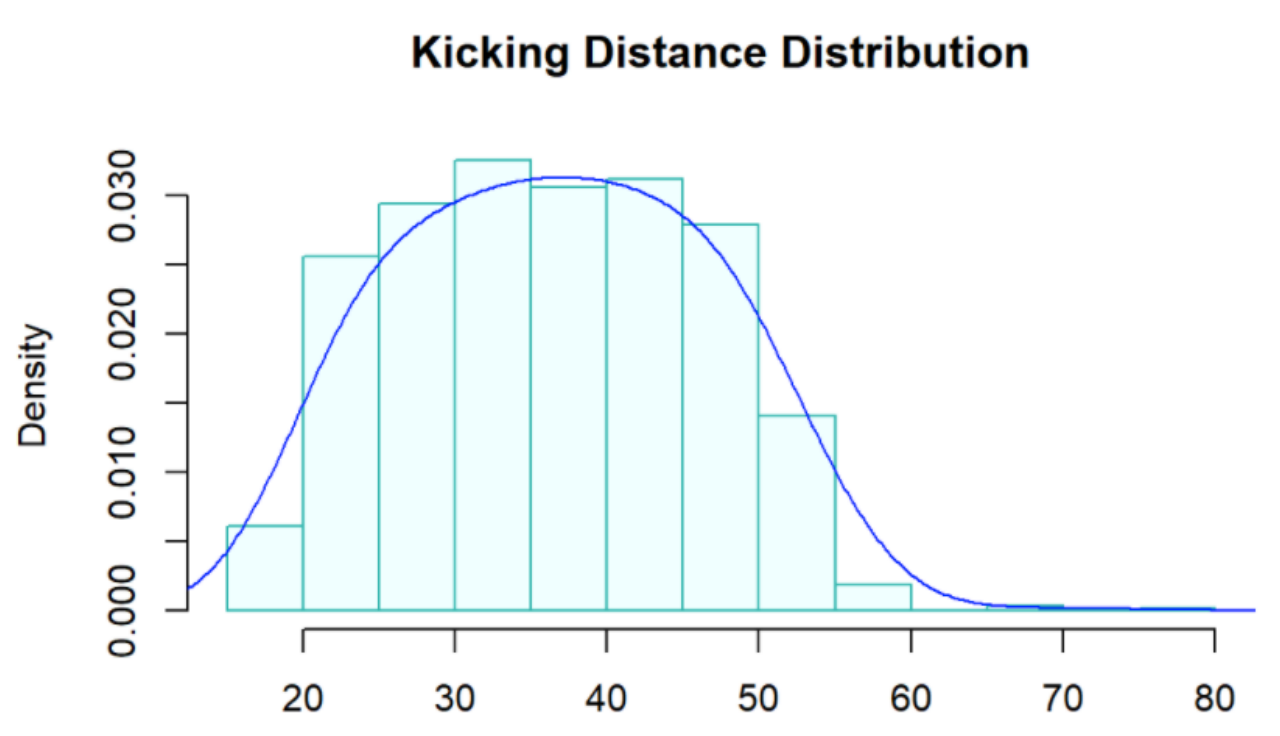
**Distance**: Numerical predictor, distance in which a football is kicked to the goal post (yards)



# RESEARCH QUESTION

- Is there any association between whether a football player can successfully make a field goal attempt and the distance that a football is kicked from during a game to score it?

# EXPLORATORY DATA ANALYSIS




- Symmetrical and unimodal, with no outliers.
- Variables had a linear relationship.
- No high intercorrelations
- There is no issue of potential imbalance.
- I did not transform the kicking distance variable.
- I fitted a logistic regression directly to the data.

# SUMMARY STATISTICS OF THE SIMPLE LOGISTIC REGRESSION MODEL COEFFICIENTS

The summary stats of regression coefficients

| | Estimate | Std. Error | z value | Pr(>\|z\|) | 2.5 % | 97.5 % |
|---|---|---|---|---|---|---|
| (Intercept) | 6.7627078 | 0.5444277 | 12.42168 | 0 | 5.7399740 | 7.877643 |
| distance | -0.1208357 | 0.0122852 | -9.83590 | 0 | -0.1457751 | -0.097540 |

- Kicking distance is negatively associated with the ability to make a field goal attempt.
- The slope of the distance variable equals -0.1208.
- As kicking distance increases by 1 yard, the chances of making a field goal attempt decrease by 0.1028.
- The p-value is 0 and the 95% confidence interval is [-0.1458, -0.0975].

# THE ODDS RATIOS

Summary Stats with Odds Ratios

| | Estimate | Std. Error | z value | Pr(>\|z\|) | odds.ratio |
|---|---|---|---|---|---|
| (Intercept) | 6.7627078 | 0.5444277 | 12.42168 | 0 | 864.9812621 |
| distance | -0.1208357 | 0.0122852 | -9.83590 | 0 | 0.8861796 |

- Odds ratio associated with kicking distance is 0.88.
- As kicking distance increases one yard, odds of making a field goal attempt decreases by about 12%.
- This is a practically significant risk factor.

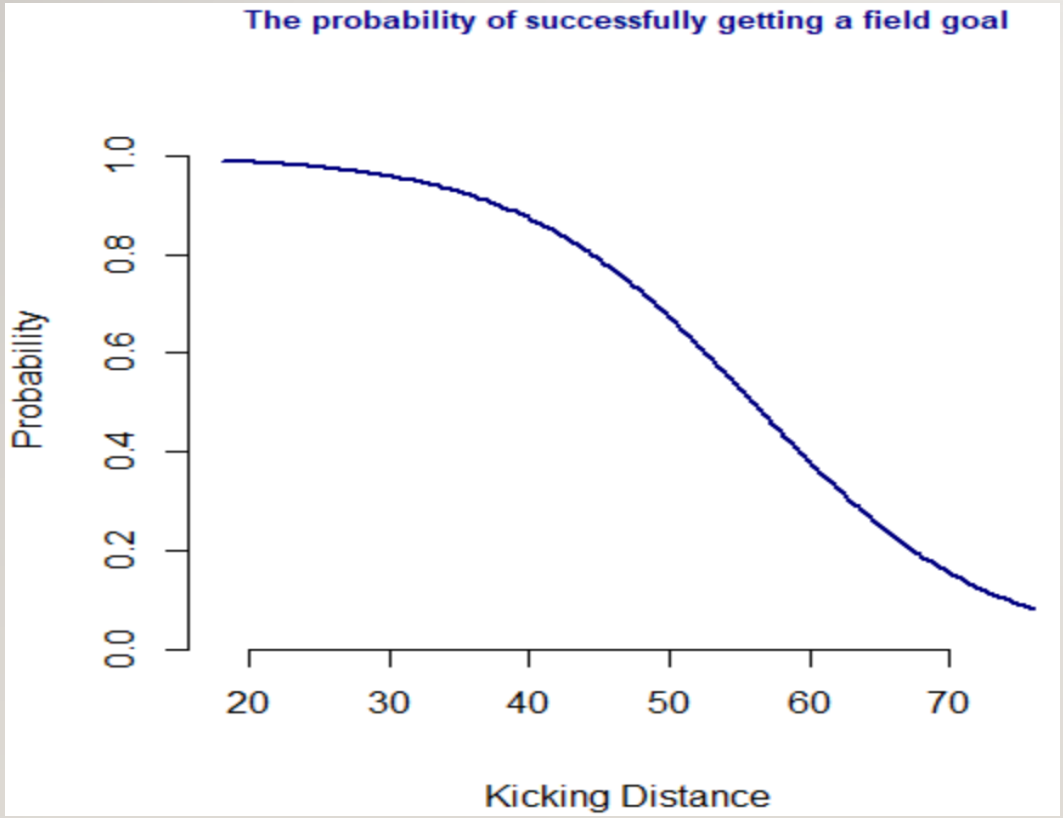
# GLOBAL GOODNESS OF FIT MEASURES

Global Goodness-of-fit Measures

| Deviance.residual | Null.Deviance.Residual | AIC |
|---|---|---|
| 686.927 | 817.7227 | 690.927 |

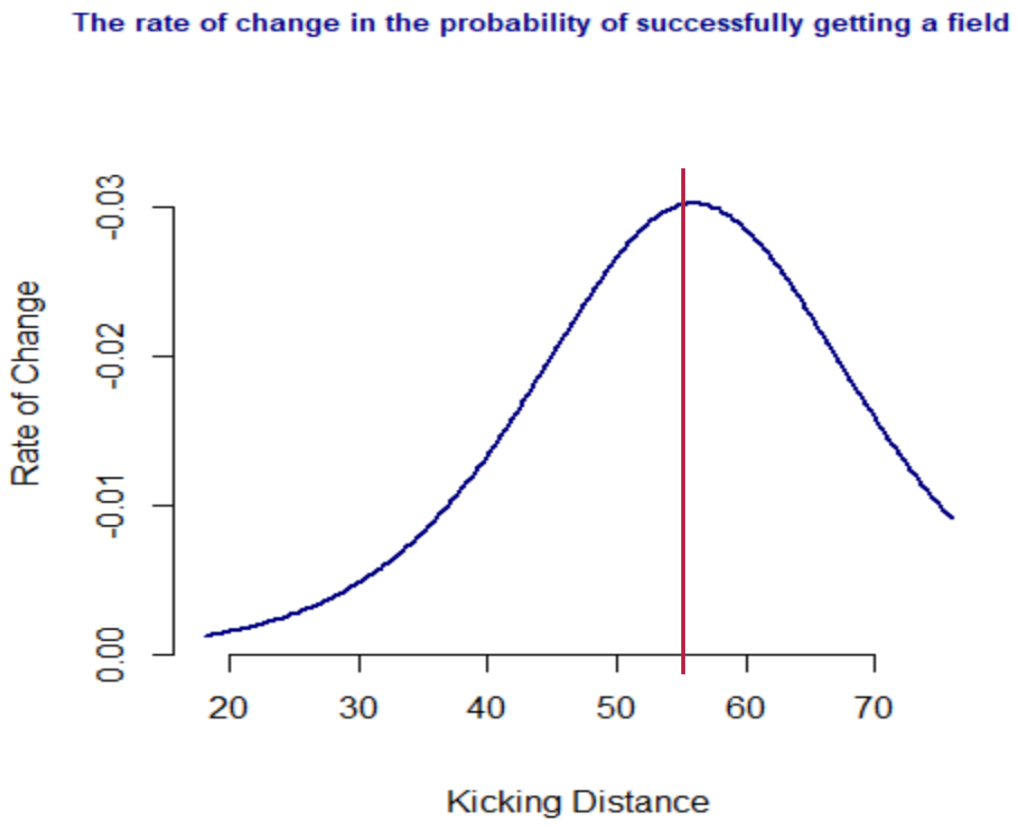
- I can't have other candidate models with corresponding likelihood at the same scale to compare.
- I will not interpret these goodness-of-fit measures.

# SUCCESS PROBABILITY CURVES




- The standard S curve represents how the probability of making a field goal attempt decreases as kicking distance increases.

# SUCCESS PROBABILITY CURVES




- This graph indicates the rate of change in the probability of making a field goal attempt decreases when kicking distance is less than 55 yards.
- The rate of change starts to increase when kicking distance is greater than 55 yards.
- The turning point is about 55 yards.

# SUMMARY

- As the distance the football is kicked from increases, the chance of making a field goal attempt decreases.
- Kicking distance is negatively associated with ability to make a field goal attempt.

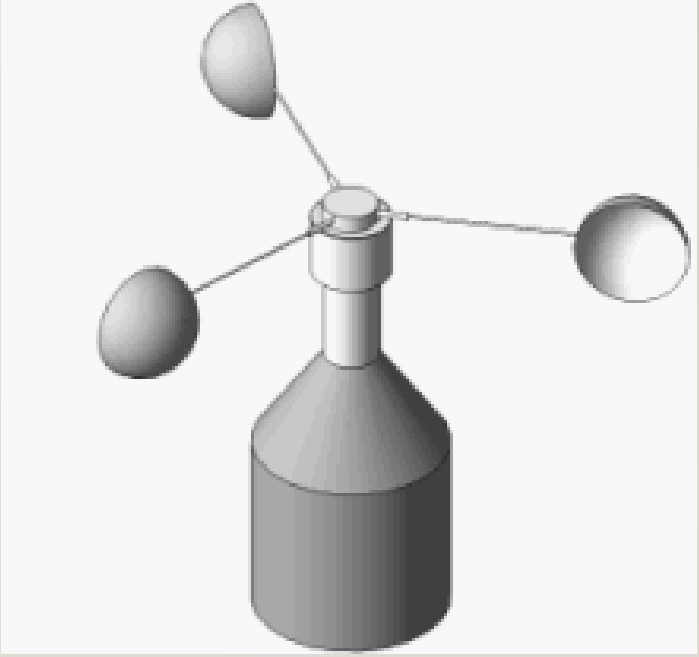
- The slope of the distance variable is -0.1208.
- As kicking distance increases by 1 yard, the chances of making a field goal attempt decrease by 0.1028.
- The p-value is 0 and the 95% confidence interval is [-0.1458, -0.0975].

- The odds ratio associated with kicking distance is 0.88.
- Since this number is less than 1, as the kicking distance increases one yard, the odds of making a field goal attempt decreases by about 12%.

- The rate of change in probability of making a field goal attempt decreases when kicking distance is under 55 yards.
- It starts to increase when kicking distance is greater than 55 yards.
- It is better for most field goal attempts to be made from at most 55 yards from the goal post.

# POSSIBLE IMPROVEMENTS

- Maybe when running this regression, the data should also be sorted by kicker.
- It would be good to include a variable for wind speed and/or direction since this affects making field goal attempts in a dramatic way.
- Data from other seasons besides 2008 should also be present for observing patterns over a wider timespan.



# SOURCES

"NFL Field Goal Attempts by Kicker with Situational Info" Data Set:

http://users.stat.ufl.edu/~winner/data/nfl2008_fga.csv