Logistic Regression Cheat Sheet

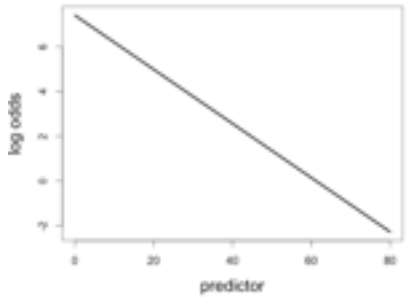
Using least squares regression makes less sense when the outcome is binary. For example, if we use a line to model the probability of an outcome, the model will predict probabilities less than zero or greater than one. Logistic regression fixes this by modeling a linear relationship between the predictors and the *log odds*.

The following three "spaces" are different ways of looking at this relationship.

Log Odds

Thinking about the log odds of the outcome is most useful when considering the linear form of the regression equation line.

$$log\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 X$$

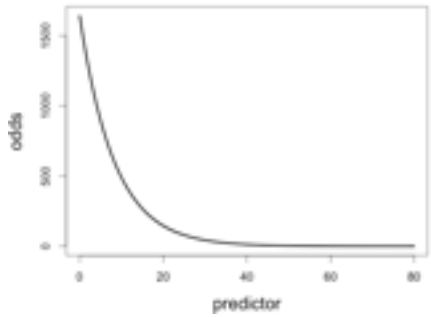


"A one unit increase in x is associated with a β_1 increase in the log odds of y."

Odds

Thinking about the odds of the outcome is most useful when considering the interpretation of the slope coefficient.

$$\frac{\pi}{1-\pi} = e^{\beta_0+\beta_1 X}$$

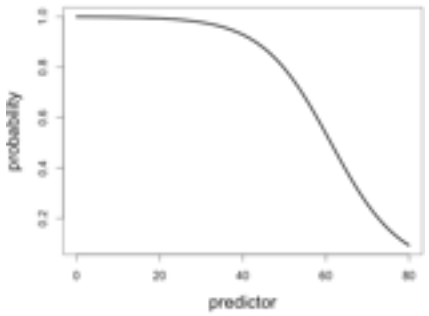


"A one unit increase in x is associated with changing y by a factor of e^{β_1}."

Probability

Thinking about the probability of the outcome is most useful when considering the model's prediction for one value of the predictor.

$$\pi = \frac{e^{\beta_0+\beta_1 X}}{1+e^{\beta_0+\beta_1 X}}$$



Note: There is no good interpretation for how a change in x is associated with a change in the probability of y.

Cheat sheet by Sara LaPlante