**Week 8, Day 1**

**Logistic Regression**

* Trashball data

trash <- matrix(nrow=6, ncol=2) #this creates a 6x2 matrix

colnames(trash) <- c("distance", "prop")

trash <- as.data.frame(trash) #turn trash into a data frame

trash$distance <- c(4, 8, 12, 16, 20, 24)

trash$prop <- c(1.0, 0.875, 0.5, 0.292, 0.083, 0.25)

plot(prop~distance,data=trash) #scatterplot

trash.lm <- lm(prop~distance,data=trash) #linear regression

abline(trash.lm)

* + Does it look like the linear model fits the data well?
  + What does the model predict for a distance of 14 feet? 1 foot? 35 feet?

🡪 Since our response variable is a proportion, we need a model that will constrain Y between 0 and 1.

* Intro to Logistic Regression
  + Logit function
  + Logit form of the model

Probability form of the model

* Odds
  + Odds = P(success)/P(failure)
  + Odds and probability are NOT the same thing!
  + Interpreting odds
* Odds ratios (ORs)
  + Odds ratio is the ratio of two odds
  + Situation 1: **Binary** explanatory variable
    - Interpreting ORs
    - Activity: Odds & ORs, Part 1