

# ML, HW 5, Due March 2

Rob McCulloch

February 22, 2017



# 1. Hockey Data Logit

Use a logistic regression to fit the hockey data from the last homework.

How does the logistic fit compare to the ones you had in the previous homework?

(to really do this would be a lot of work, you can just make qualitative comparisons).

Take the goaldiff out of the model and compare the AIC and BIC with and without goaldiff. What do AIC and BIC say about goaldiff?

## 2. Satisfaction Data

Get the data “satisfaction.csv” from the webpage.

The “y” is how highly residents rate their town.

The basic version of y is the variable yr.

yrb is a binary version of y and yr3 is a version of y with three categories.

There are two x variables: gender and age (age is already scaled to (0,1)).

Use logistic regression to relate yrb to gender and age.

Fit with and without gender and age.

Which model is the best according to AIC?

Which model is the best according to BIC?

Use multinomial regression to relate yr3 to gender and age.

Fit with and without gender and age.

Which model is the best according to AIC?

Which model is the best according to BIC?

Can you interpret the coefficients for age?

Is an older person more or less likely to like his town???