Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Worksheet 7

**1)** You have a dataset with 4 columns, the first several rows of which are shown below:



This file was read into R with the following line of code:

dat <- read.csv(“grasp.csv”)

Write the code to fit a GLM to this data where the first column is the response variable and columns 2-4 are all included predictor variables.

**2)** You are studying fighting behavior among a group of colonial nesting birds. There is a hypothesis that when birds lose a fight, they are more likely to start a fight to regain social standing. You decide to test this by asking whether or not birds that lose a fight are more likely to initiate a subsequent fight on the same day than you would expect by chance. In your analysis of the empirical data, of the 87 fight losers in your dataset, 23% initiated a subsequent fight later that same day. In your Monte Carlo, 378 of the 10000 simulations had 20 or more fight losers that initiated a subsequent fight later that same day. How do you interpret this result; what is the p-value for this test?

**3)** What are some of the challenges that you face when you attempt to perform a GWAS study?

**4)** When are Monte Carlo methods a good choice for testing a hypothesis?