Exercise 9.1

A survey of a random sample of students at the University of New Hampshire was conducted. We are interested in predictors of aggression. Students were classified as either having or not having aggression (`aggression`: 1=has aggression, 0=does not) based on responses to a validated questionnaire. We are interested in potential differences in aggression comparing students who belong and don't belong to a fraternity/sorority (`frat`: 1=belong to a fraternity or sorority, 0=do not belong to one).

(a) A logistic regression model was run using fraternity/sorority membership to predict aggression (output below). Based on this model, interpret the effect of fraternity/sorority membership as an odds ratio.

```
. logistic aggression frat, coef

Logistic regression                             Number of obs     =        243
                                                LR chi2(1)        =       6.78
                                                Prob > chi2       =     0.0092
Log likelihood = -161.90393                     Pseudo R2         =     0.0205

------------------------------------------------------------------------------
  aggression |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        frat |    .906812   .3639366     2.49   0.013     .1935094    1.620115
       _cons |   .1636294   .1433355     1.14   0.254     -.117303    .4445619
------------------------------------------------------------------------------
```

(b) Is there a significant effect of fraternity/sorority membership on aggression? Cite specific output in your answer. (Assume $\alpha = 0.05$)

(c) There is a 95% CI for `frat` on the logistic regression output. Report this CI, state what population parameter this is a CI for, and explain how the 95% CI could be used to answer part (b).

(d) Calculate the estimated probability of aggression for people not in a fraternity/sorority based on the logistic regression result.

(e) Next, belonging to a fraternity/sorority and aggression were cross-tabulated (output below). Use this to calculate the appropriate odds ratio to confirm your answer to (a) – the estimated OR from the logistic regression.

```
. tabulate aggression frat

           |         frat
aggression |         0          1 |     Total
-----------+----------------------+----------
         0 |        90         12 |       102
         1 |       106         35 |       141
-----------+----------------------+----------
     Total |       196         47 |       243
```

(f) Use this cross-tabulation to confirm your answer to (d) – the estimated probability of aggression among people not in a fraternity/sorority.

Exercise 9.2

Using the New Hampshire survey, we are also interested in whether age is a predictor of aggression. A logistic regression model was run using age (age, in years) to predict aggression (aggression: 1=has aggression, 0=does not). Note that the coef option was left off the logistic command.

```
. logistic aggression age

Logistic regression                               Number of obs     =        243
                                                  LR chi2(1)        =       3.88
                                                  Prob > chi2       =     0.0489
Log likelihood = -163.35176                       Pseudo R2         =     0.0117

------------------------------------------------------------------------------
  aggression | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .9132199   .0453706    -1.83   0.068      .8284875    1.006618
       _cons |   8.774296   8.920195     2.14   0.033      1.196346    64.35283
------------------------------------------------------------------------------
```

(a) Interpret the effect of age on aggression. Is there a significant effect? (Assume $\alpha = 0.05$)

(b) There is a 95% CI for age on the logistic regression output. Report this CI, state what population parameter this is a CI for, and explain how the 95% CI could be used to determine if there is a significant effect of age on aggression.

Exercise 9.3

Using the New Hampshire survey, we are also interested in whether there are differences in the probability of having aggression across the years in school. A logistic regression model was run using year in school (`year`, 1=freshman, 2=sophomore, 3=junior, 4=senior) to predict aggression (`aggression`: 1=has aggression, 0=does not). Use the Stata output on the next page to answer the questions below.

(a) Is there evidence of significant differences in aggression by year in school? Provide specific evidence to support your claim. (Assume $\alpha = 0.05$)

(b) Interpret completely the effect of year in school on aggression (using odds ratios). For each comparison you interpret, provide evidence that the comparison is or is not significant. (Assume $\alpha = 0.05$)

(c) CHALLENGE QUESTION: What is the estimated odds ratio comparing the odds of aggression for juniors compared to seniors?

```
. generate year2 = (year==2) if !missing(year)
. generate year3 = (year==3) if !missing(year)
. generate year4 = (year==4) if !missing(year)
. logistic aggression year2 year3 year4, coef

Logistic regression                               Number of obs     =        243
                                                  LR chi2(3)        =      10.16
                                                  Prob > chi2       =     0.0173
Log likelihood = -160.21221                       Pseudo R2         =     0.0307

------------------------------------------------------------------------------
  aggression |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       year2 |   -.040822   .4143268    -0.10   0.922    -.8528875    .7712435
       year3 |   .1823216   .4082483     0.45   0.655    -.6178304    .9824735
       year4 |   -.863647   .4149159    -2.08   0.037    -1.676867   -.0504267
       _cons |   .5108256   .3265986     1.56   0.118    -.1292959    1.150947
------------------------------------------------------------------------------

. test year2 year3 year4

 ( 1)  [aggression]year2 = 0
 ( 2)  [aggression]year3 = 0
 ( 3)  [aggression]year4 = 0

           chi2(  3) =    9.92
         Prob > chi2 =    0.0193
```