Criminal Justice Doctoral Comprehensive Exam
Statistics
August 2016

There are two questions on this exam. Be sure to answer ***both*** questions in the 3 and half hours to complete this exam. Read the instructions carefully to be sure you are answering the correct part(s) of each question. Each question has associated output, be sure to refer to the output listed in each question.

Question 1:

We are interested in addressing two hypotheses around self-reported fear of victimization. A random survey of 200 respondents were asked whether or not they were afraid of being victimized (0=No, 1=Yes). They were further asked of the following three locations which did they feel presented the most potential risk for victimization (1=Home, 2=Outside, 3=Work).

The independent variables included were:

- Age (continuous)
- Gender (0=Male, 1=Female)
- SES (1=Low, 2=Medium, 3=High)
- Prior Victimization (0= Not Victimized, 1=Victimized)

<u>For each of the two sets of output (Part A=Logistic; Part B=Multinomial) address the following:</u>

1. Identify the appropriate research and null hypotheses.
2. What do we know about model fit, in regards to the results presented here? Are there alternatives that might be preferable?
3. Explain why the logistic or multinomial logistic model was the most appropriate statistical test to utilize.
4. Interpret the significant coefficients. Be as specific as possible.
5. What conclusions do you draw about your original hypotheses?
6. Consider the two alternative types of statistics utilized here. Is there a way that one could be used to inform the other?

Question 2:

In the output for question 2 you will see the results of a linear regression analysis that asked participants about their confidence in local law enforcement (Conf-PO- higher values indicate greater confidence). This is the same database as used for question 1. With these results, answer the following questions:

1. What do we know about the distribution of the dependent variable, if we are confident that a linear regression is the most appropriate statistical test to use?

2. What do we know about the model fit, as presented here?

3. Interpret all of the variable coefficients in the model.

4. What if the dependent variable had not met our criteria to conduct a linear regression analysis? What would our alternatives have been?

5. Based on what you know about these variables from question 1, why could we not include the dichotomous measure of fear in the model?

6. Give two examples of statistical analyses that you would likely have conducted prior to running this linear regression. In other words, what analyses would you have done in order to be confident that a linear regression was appropriate? Be sure to explain your answers.

Output for Question 1, Part A

```
. logit Fear Sex Age Prior i.SES, or

Iteration 0:   log likelihood = -138.58943
Iteration 1:   log likelihood = -114.10718
Iteration 2:   log likelihood = -113.78446
Iteration 3:   log likelihood = -113.78395
Iteration 4:   log likelihood = -113.78395

Logistic regression                             Number of obs     =        200
                                                LR chi2(5)        =      49.61
                                                Prob > chi2       =     0.0000
Log likelihood = -113.78395                     Pseudo R2         =     0.1790

------------------------------------------------------------------------------
        Fear | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         Sex |   3.023224   1.031464     3.24   0.001     1.549021    5.900426
         Age |   1.068578   .0467533     1.52   0.130     .9807622    1.164257
       Prior |   3.337546   1.891575     2.13   0.033     1.099015    10.13563
             |
         SES |
      Medium |   1.480597   .5921421     0.98   0.326     .6761016    3.242365
        High |   6.186664   2.449851     4.60   0.000     2.847018    13.44383
             |
       _cons |   .0266728   .0371567    -2.60   0.009      .001739    .4091108
------------------------------------------------------------------------------
```

Output for Question 1, Part B

```
. mlogit Fear_Loc Sex Age Prior i.SES, b(2) rrr

Iteration 0:   log likelihood = -198.65125
Iteration 1:   log likelihood = -192.25956
Iteration 2:   log likelihood =  -192.0156
Iteration 3:   log likelihood = -192.01446
Iteration 4:   log likelihood = -192.01446

Multinomial logistic regression                 Number of obs     =        200
                                                LR chi2(10)       =      13.27
                                                Prob > chi2       =     0.2088
Log likelihood = -192.01446                     Pseudo R2         =     0.0334

------------------------------------------------------------------------------
    Fear_Loc |        RRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Home         |
         Sex |   .7961732   .3006814    -2.60   0.002     .3797906    .869056
         Age |   1.015737   .0502956     0.32   0.753     .9217919    1.119257
       Prior |   .2108194   .6223345    -3.37   0.000     .1421607    .315723
             |
         SES |
     Medium  |    2.16275   1.022553     1.63   0.103     .8561625    5.463315
       High  |   1.186343   .5637197     0.36   0.719     .4674566    3.010783
             |
       _cons |   .1820584   .2808919    -1.10   0.270     .0088495     3.74543
-------------+----------------------------------------------------------------
Outside      |  (base outcome)
-------------+----------------------------------------------------------------
Work         |
         Sex |   1.323817   .4880723     0.76   0.447     .6426912    2.726803
         Age |   .9547572   .0443528    -1.00   0.319     .8716672    1.045768
       Prior |   .2340993   .1829387    -1.98   0.047     .0506083    .965327
             |
         SES |
     Medium  |   1.837762   .8558437     1.31   0.191     .7377098     4.57818
       High  |   .8297222   .8675656     2.66   0.018     .6782271    .991009
             |
       _cons |   1.101233   1.570022     0.07   0.946     .0673482    18.00663
------------------------------------------------------------------------------
```

Output for Question 2

```
. regress Conf_PO Sex Age Prior i.SES

      Source |       SS           df       MS      Number of obs   =       200
-------------+----------------------------------   F(5, 194)       =     11.62
       Model |  6256.36405         5  1251.27281   Prob > F        =    0.0000
    Residual |  20885.0559       194  107.654928   R-squared       =    0.2305
-------------+----------------------------------   Adj R-squared   =    0.2107
       Total |    27141.42       199  136.389045   Root MSE        =    10.376

------------------------------------------------------------------------------
     Conf_PO |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         Sex |  -.4248053   1.509475    -0.28   0.779    -3.401895    2.552284
         Age |   .4664033   .1944315     2.40   0.017     .0829323    .8498743
       Prior |  -16.44588   2.260582    -7.28   0.000    -20.90435   -11.9874
             |
         SES |
      Medium |   2.063475   1.903511     1.08   0.280    -1.690758    5.817709
        High |   1.467059    1.78326     3.26   0.001     1.050007    3.984125
             |
       _cons |   46.36742   6.012069     7.71   0.000     34.51001    58.22483
------------------------------------------------------------------------------
```