**Week 9 Assignment**

Michael K. Zoucha

College of Science & Information Technology, Bellevue University

DSC 510: Introduction to Programming

Professor Eller

25 October 2021

**Week 9 Assignment**

Exercise 11-1:

Import necessary packages and create data frame with pregnancies greater than 30 weeks along

import first  
live, firsts, others = first.MakeFrames()  
live = live[live.prglngth>30]

Use ordinary least squares, get p-value to determine statistically significant effect of different predictor variables

import statsmodels.formula.api as smf  
model = smf.ols('prglngth ~ birthord + npostsmk + workpreg + agecon', data=live)  
results = model.fit()  
print(results.summary())

With a p-value of less than 0.05, we can determine that working while pregnant has an statistically significant effect on pregnancy length, but smoking and mother’s age do not.

A picture containing text

Description automatically generated

Exercise 11-3:

Add age\_squared column to the data frame for analysis, then use age, age\_squared, race, totinc, and educat for Poisson regression analysis. (count)

join['age2'] = join.age\_r\*\*2  
formula='numbabes ~ age\_r + age2 + C(race) + totincr + educat'  
model = smf.poisson(formula, data=join)  
results = model.fit()  
print(results.summary())

We can see all included predictor variables are statistically significant by the very low p-values.  
Graphical user interface, text

Description automatically generated with medium confidence

Create a new data frame with the values to predict how many kids she will have

columns = ['age\_r', 'age2', 'race', 'totincr', 'educat']  
new = pd.DataFrame([[35, 35\*\*2, 1, 14, 16]], columns=columns)  
results.predict(new)

Results:  


Exercise 11-4:

Run model to guess status of marriage using age, age\_squared, race, totincr, and educat:

formula='rmarital ~ age\_r + age2 + C(race) + totincr + educat'  
model = smf.mnlogit(formula, data=join)  
results = model.fit()  
print(results.summary())

Table

Description automatically generated with medium confidenceTable

Description automatically generated

Create new frame to predict marriage status:

columns = ['age\_r', 'age2', 'race', 'totincr', 'educat']  
new = pd.DataFrame([[25, 25\*\*2, 2, 11, 12]], columns=columns)  
print(results.predict(new))

Results:

There is a 75% chance she is married.

