BUS 735: Business Decision Making and Research
Homework: Logistic Regression

Learning Objectives:

- LO2: Be able to construct and use multiple regression models (including some limited dependent variable models) to construct and test hypotheses considering complex relationships among multiple variables.
- LO6: Be able to use standard computer packages such as R to conduct the quantitative analyses described in the learning objectives above.
- LO7: Have a sound familiarity of various statistical and quantitative methods in order to be able to approach a business decision problem and be able to select appropriate methods to answer the question.

Directions: Type up your answers in a single word document, and include the relevant R codes and output that you cite, copied and pasted into the word processing document. When asked "Test the hypothesis.." or "Is there evidence of.." or "Is there statistical significance of.." or "Is there a relationship of..." conduct the appropriate hypothesis test, following these steps (in order):

- Indicate what statistical test / statistical method you are using.
- State the null and alternative hypothesis.
- Report the p-value.
- Conclude whether you reject or fail to reject the null hypothesis.
- State your result in plain English.

The homework assignment uses a data set consisting of 753 married women in the United States and information about whether they participate in the labor market (either they have a job or are actively looking for one) and background information on them and their families.

The data can be downloaded from one of the following sources:

- http://murraylax.org/datasets/mroz.RData
  (download then use command `load(mroz.RData')`)
- http://murraylax.org/datasets/mroz.csv
  (use command `dat <- read.csv('http://murraylax.org/datasets/mroz.csv)`)

The variables in the data set that are used in this assignment include:

- `inlf`: Short for "In Labor Force" which is a dummy variable equal to 1 if the woman is in the labor force and 0 if not. In the labor force means that the woman is either employed or is actively seeking employment.
- `kidslt6`: Number of children under age of 6.
- `kidsge6`: Number of children age 6-18.
- `age`: age of the woman

- `educ`: Number of years of education of the woman
- `hushrs`: Number of hours per year that the husband works.
- `huseduc`: Number of years of education of the husband.
- `amotheduc`: Number of years of education of the woman's mother.
- `fatheduc`: Number of years of education of the woman's father.
- `city`: Dummy variable equal to 1 if the woman lives in a city (metropolitan statistical area).

1. Is there a relationship between the years of education of a women and whether or not she decides to participate in the labor market? If so, do women who participate in the labor market on average have more or less years of schooling that women who do not?
2. Is there a relationship between the years of education of the woman and the years of education of the husband? If so, describe the nature of the relationship.
3. Is there a relationship between whether or not the woman participates in the labor force and the number of children under the age of 6.
4. Estimate a logistic regression using `inlf` as the outcome variable and all other variables as explanatory variables, and answer the following questions:
   (a) For which variables is there statistical evidence that the variable influences whether or not the woman participates in the labor market? For the explanatory variables where there is statistical evidence, comment on the direction of influence on the probability of being in the labor force.
   (b) What is the marginal effect of number of children under 6 years on the probability for an average woman being in the labor force?
   (c) What is the probability that a woman with the following characteristics will participate in the labor market?
      - The woman has three kids, ages 4, 6 and 7.
      - The woman is 38 years old.
      - The woman has 16 years of education.
      - The husband works 2500 hours per year.
      - The husband has 12 years of education.
      - The woman's father and mother each have 16 years of education.
      - The woman does lives in a city.

   (d) How much more likely does a woman like the one described in the problem above participate in the labor market if she has one additional year of education?
   (e) How much more likely does a woman like the one described in the problem above participate in the labor market if she has one additional child over the age of 6?