BUS 735: Business Decision Making and Research
Instructor: Dr. James Murray
Fall 2012
Take Home Exam 1

Directions: Type up answers to all of the following questions. Include in your document only the relevant SPSS output that you need to answer the question. Please copy and paste this SPSS output; do not include in your submission any other files except a single Microsoft Word document or PDF document that includes all your answers with the relevant SPSS output accompanying each answer. Every time you conduct a hypothesis test, indicate what statistical test you are using, what are the null and alternative hypotheses, what is your p-value, and a plain English description of what is your conclusion. **Exam is due on Wednesday, October 31 by 5:29pm**.

1. The dataset `salesrev.xls` contains the monthly sales revenue (in thousands of dollars) for 200 sales people for a large national corporation. The sales people focus on one of two categories of products, which are labeled as Product 1 and Product 2 in the dataset. Also included in the dataset is the years of experience each sales person has. The company introduces a new sales person training program. The training program involves three full day training sessions, one training session is offered each month for three months. The 200 sales people participate in the training program, and the company gathers data on the monthly sales before the training (`Sales0`), after the first training session (`Sales1`), after the second training session (`Sales2`), and after the third and final training session (`Sales3`). For the following questions, test the appropriate hypothesis and report your conclusion.

   (a) Is there evidence that average sales is different between the times when the sales people have not yet had training to after they have completed all their training?

   (b) Is there evidence that average sales are different between those that sell product 1 versus product 2?

   (c) Taking into account the effect the product type and years experience has on sales revenue, is there evidence that the training program positive influences sales?

   (d) Based on the results above, do all of the training sessions lead to an increase in sales revenue? If not, which training sessions do you find evidence that are effective, and which training sessions do you fail to find evidence that they are effective?

   (e) Taking into account the effect of training and years experience, is there a difference in sales revenue between sales people who sell Product 1 versus Product 2?

   (f) Based on the results above, are the training sessions more effective for sales of either product 1 or product 2?

2. The dataset `wage1D.xls` contains the following variables including wage and background information for 526 individuals:

- `wage`: average hourly earnings
- `educ`: years of education
- `exper`: years of experience
- `tenure`: years with current employer
- `nonwhite`: Dummy variable = 1 if employee is non-white.
- `female`: Dummy variable = 1 if employee is female.

(a) Estimate a regression that explains average hourly earnings using all the variables in the dataset. What is your estimated regression equation?

(b) Test the hypothesis that at least one of your explanatory variables helps explain average hourly earnings.

(c) Report and interpret the value for R^2.

(d) Accounting for the other variables in the model, is there evidence that non-white employees receive lower wages on average than white employees?

(e) Accounting for the other variables in the model, is there evidence that female employees receive different wages on average than male employees?

(f) What is the predicted wage for a white female with 12 years of education, 4 years of experience, 2 years with their current employer?

3. The dataset `cps78_85.xls` contains information about union membership and background characteristics for 1084 individuals. The variables include,

- `educ`: years of education
- `south`: dummy variable = 1 if employee lives in the South
- `nonwhite`: dummy variable = 1 if employee is not white
- `female`: dummy variable = 1 if employee is female
- `exper`: years of experience
- `y85`: dummy variable = 1 if year of the observation is 1985, = 0 if the year of the observation is 1978
- `union`: dummy variable = 1 if the employee is a member of a labor union.

(a) Estimate a regression that predicts the probability that a person is a member of a union based on all the other variables given above. What is your estimated regression equation?

(b) Is there evidence that males and females have different propensities to be a member of a union, given the other variables in your model? If so, which gender is more likely to be a member of a union?

(c) Is there evidence that the propensity to be a member of a union was different in 1985 versus 1978? If so, in which year was union membership more popular?

(d) What percentage of the sample who were union members did your regression model correctly predict would be union members?

(e) What percentage of the sample who were not union members did your regression model correctly predict would not be union members? Based on this and your previous answer, comment on the accuracy of your regression model.

(f) Use your regression model to predict the probability that a white man from Wisconsin with 12 years of education and 4 years of experience was a member of a union in 1978?

(g) What is the marginal effect on the probability of union membership for living in the South, for a person similar to the one described in the previous question?