BUS 735: Business Decision Making and Research
Instructor: Dr. James Murray
Take Home Exam 1 - Fall 2014
Due Monday, October 27, 2014

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.

1. The dataset `wage1D.sav` 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) What percentage of the variability in average hourly earnings is explained by education, experience with employer, total experience, race, and gender?

   (c) Controlling for the other variables in the model, is there evidence that female employees receive lower wages on average than male employees? If so, what is the estimated difference in hourly earnings?

   (d) Controlling for the other variables in the model, is there evidence that more total experience leads to an increase in average hourly earnings? If so, what is the estimated difference in average hourly earnings between someone who has 8 years of experience, and someone who has 10 years of total experience?

   (e) What is the predicted wage for a white male with 12 years of education, 10 years of experience, and 8 year experience with his current employer?

2. The following questions use the dataset `cex.sav`. This is recent (2010:Q2) consumer income data from the Current Population Survey. The variables included in this SPSS file include:

- Age (in years)
- Relationship to head-of-household: 1=head of household, 2=spouse, 3=child or adopted child of head, 4=grandchild of head, 5=in-law of head, 6=brother/sister of head, 7=mother/father of head, 8=other relatives, 9=unrelated individual, 0=na.
- Education: 00=Never attended school, 1-11 1st grade through 11th grade, 38=Twelfth grade no degree, 39=High school graduate, 40=Some college no degree, 41=Associate's degree (occupational/vocational), 42=Associate's degree (academic), 43=Bachelor's degree, 44=Master's degree, 45 =Professional degree, 46=Doctorate degree
- Race: 1=White, 2=Black/African American, 3=American Indian or Aleut Eskimo, 4=Asian or Pacific Islander, 5=other
- Gender: 1=male, 2=female
- Marital Status: 1=married, 2=widowed, 3=divorced, 4=separated, 5=never married
- Employee Status: 1=member worked full time for a full year, 2=member worked part time for a full year, 3=member worked full time for part of year, 4=member worked part time for part of year, blank if member did not work.
- Employee Type: 1=private company, 2=government employee, 3=self-employed, 4=working without pay.
- Hours worked per week
- Weeks worked per year
- Occupation: 01=managerial and professional specialty occupation, 02=technical, sales, and administrative support occupations, 03=service occupations, 04=farming, forestry, and fishing occupations, 05=precision production, craft, and repair occupations, 06=operators, fabricators, and laborers, 07=armed forces, 08=self employed, 09=not working, 10=retired, 11=other, including not reported.
- Total Income: in dollars.

Transform Education into a new variable with the following categories:

(a) =1 if High School graduate or less
(b) =2 if Any college below a Bachelor's degree (40<EDU<43)
(c) =3 if Bachelor degree
(d) =4 if Masters, professional, or doctorate degree

For all of the following questions, use the results from a single analysis that controls for how all of the following variables affect income: employee type, gender, race, age, hours worked per week, and education level (as defined by your new variable).

(a) Comment on all the factors for which you find statistical evidence that influence average income. Report explicitly your hypothesis tests for these variables.
(b) For the categorical variables that you identify above as influencing average income, comment on the relationship between the factor and average income. What categories of individuals do you find statistical evidence for having higher income?
(c) For the interval/ratio variables that you identify above as influencing average income, comment on the relationship between the factor and average income.
(d) For what *pairs* of categorical factors do you find statistical evidence for an interaction effect? Conduct these hypothesis tests and comment on the nature of the interaction effect.

3. The dataset `LoanApplications.sav` contains data collected as part of the Housing Mortgage Disclosure Act on all mortgage loans in the state of Wisconsin for 2013 for first-mortgages for owner occupied housing (more than 58,000 observations). The variables include the following:

- `LoanAmt`: Amount of the loan request (in thousands of dollars)
- `PreapproveRequest`: Whether or not pre-approval was requested (=1 if requested, =2 if not, and =3 if not applicable)
- `ActionType`: Whether the loan was approved by the lending institution and accepted by the borrower (=1 if approved and accepted, =2 if approved but not accepted, =3 if denied).
- `County`: County in Wisconsin, number instead of descriptive text.
- `Ethinicity`: Ethnicity of the applicant (first applicant in the case of co-applicants). =1 if non-Hispanic / non-Latino/a, =2 if Hispanic or Latina
- `Race`: Race of the applicant (first applicant in the case of co-applicants). =1 if White, =2 Black / African American, ...
- `Sex`: Sex of the applicant (first applicant in the case of co-applicants). =1 if Male, =2 if Female
- `Income`: Combined household income of the applicants and co-applicants (in thousands of dollars)

(a) Not controlling for any other variables, is there a relationship between the loan amount and the income of the applicant(s)? If so, describe the relationship.

(b) Not controlling for any other variables, is there a relationship between whether pre-approval was sought on the loan and the subsequent action taken on the loan.

(c) Not controlling for any other variables, is there a relationship between race and the action taken on the loan?

(d) Not controlling for any other variables, is there a difference in the proportion of men approved for a loan versus the proportion of women approved?

4. Use the same dataset as the previous question, and answer the following questions by estimating a logistic regression that predicts whether or not an applicant is approved for a loan, based sex, ethnicity, whether or not the person is white, and the ratio of amount of their loan request to their income.

(a) Write down the estimated logistic regression equation.

(b) Accounting for the controls in your logistic regression, is there evidence that sex influences whether or not someone will be approved for a loan?

(c) Accounting for the controls in your logistic regression, is there evidence that race influences whether or not someone will be approved for a loan?

(d) Predict the probability that the following person will be approved for a loan. The person is male, white, non-Hispanic / non-Latino, and the ratio of their loan request to their income is 2.5.

(e) How much more or less likely would the person above be approved for a loan if he was not white?