AGEC5213: *ECONOMETRIC METHODS*
Spring 2019

PROBLEM SET NO. 1
Date Due: February 13, 2019

I. Estimate, Standard Error, t-statistic, and R-squared (5 points)

Researchers are often concerned with how consumers adjust their expenditures on some of addictive commodities (e.g., tobacco, alcohol) as their level of income increases. Consider the following econometric model that relates the proportion a household's income spent on alcohol (ALCOHOL), household's income (INCOME), age of the household head (AGE), and the number of children in the household (NUMKID).

$$ALCOHOL = \beta_1 + \beta_2 \ln(INCOME) + \beta_3 AGE + \beta_4 NUMKID + e$$

A household survey data were used to estimate this model, and the SAS output is kindly presented below. So you don't need to estimate this model by yourself (you can take a nap for a while). One minor problem is that my old laser printer has a problem with its ink cartridge, and as a result some parts of the print-out are left blank.

$$t = \frac{b_K - \beta_K}{se(b_K)}$$

Dependent variable: ALCOHOL
Included Observation: 1519

Variable	Coefficient b_K	Std. Error $se(b_K)$	t-Statistic (t)	Prob. β_K
C	0.009052	0.024050	(i)	0.7347
Ln(INCOME)	0.527641	(ii)	6.608620	0.0000
AGE	(iii)	0.00208	-6.962389	0.0000
NUMKID	-0.013282	0.003259	-4.074993	0.0000
R-squared	(iv)	Mean dependent var.		0.060596
Adjusted R-squared	0.053047	S.D. dependent var.		0.063325
S.E. of Regression	(v)			
Sum squared error	8.752896			

(a) Fill in the blank spaces that appear in the output.
(b) Interpret each of the estimates b_2, b_3, and b_4.
(c) Compute 95% interval estimates for β_2 and β_3. What do these interval estimates tell you?
(d) Test the hypothesis that the household income proportion for alcohol does not depend on the number of children in the household and interpret the test result.

II. Estimates, Standard Errors, and Hypothesis Test (7 points)

Consider the regression model:

$$y_t = \beta_1 + \beta_2 x_{2t} + \beta_3 x_{3t} + e_t; \quad e_t \sim N(0, \sigma^2)$$

Data on this three-variable regression model yields the following results:

$$X'X = \begin{bmatrix} 25 & 0 & 0 \\ 0 & 10 & 20 \\ 0 & 20 & 50 \end{bmatrix}; \; X'y \begin{bmatrix} 100 \\ 1 \\ 8 \end{bmatrix}; \; \sum_t y_t^2 = 450 \qquad T=25$$

(a) What is the sample size? 25
(b) Estimate parameters of the regression model.
(c) Estimate standard errors of b_2 and b_3.
(d) Test the hypothesis that β_2=0 against $\beta_2 \neq 0$; and the hypothesis that β_3= 1 against $\beta_2 \neq 1$ (use 5% significance level for both tests).
(e) Compute R^2 and adjusted R^2 (both R^2 should be mean-centered).

III. Elasticity, R-Squared, Adjusted R-Squared, and Hypothesis Test (8 points)

Data on beginning salary (*Y*) and education in years (*X*) for 93 employees of Stillwater Bank can be found in the file, HW1-DATA.xls (see our D2L site). Estimate a log-log linear relationship between salary and education, i.e.,

$$lnY_t = \beta_1 + \beta_2 lnX_t + e_t.$$

(a) Does the coefficient of education have the expected sign? What interpretation can you place on the slope coefficient?

(b) Predict the starting salary of an individual with 13 years of education.

(c) What is the interpretation of R-square from your output?

(d) Compute adjusted R-square and provide justification of using the adjusted R-square.

(e) Test if the elasticity of salary w.r.t. education is one. Test if all coefficients are zero. List null and alternative hypotheses, the choice of probability distribution, and degree of freedom for these tests.

(f) In addition to observations on *Y* and *X*, the data contains observations on number of months of previous work experience (*E*) and the number of months that the individual was hired for the Stillwater Bank (*T*). Run the regression using the same functional form, a double log form. Did b_1, b_2, R-squared, adjusted R-squared changed? What can you say about these different regression results?

(g) Test if elasticities of the number of months of previous work experience (*E*) and the number of months that the individual was hired for the Stillwater Bank (*T*) are the same. List null and alternative hypotheses, the choice of probability distribution, and degree of freedom for this test.