Econ 3040 – Midterm, Oct. 22nd, 2015

You may use a calculator. Answer all questions in the answer book provided. The exam is 75 minutes long and consists of 72 marks. (Approximately 1 mark per minute).

A formula sheet, and a table of probabilities from the standard Normal distribution, are provided at the back of the exam booklet.

NAME:	
STUDENT #:	

DO NOT OPEN THIS BOOKLET UNTIL INSTRUCTED TO DO SO.

HAND IN THIS BOOKLET AT THE END OF THE EXAM.

Part A – Multiple Choice – 3 marks each

1.) An estimator is

a. an estimate.
b. a formula that gives an efficient guess of the true population value.
c. a random variable.
d. a nonrandom number.

2.) The correlation coefficient

a. lies between zero and one.
b. is a measure of linear association.
c. is close to one if *X* causes *Y*.
d. takes on a high value if you have a strong nonlinear relationship.

3.) When the estimated slope coefficient in the simple regression model, $\hat{\beta}_1$, is zero, then

a. $R^2 = \overline{Y}$.
b. $0 < R^2 < 1$.
c. $R^2 = 0$.
d. $R^2 > (SSR/TSS)$.

4.) The slope estimator, $\hat{\beta}_1$, has a smaller standard error, other things equal, if

a. there is more variation in the explanatory variable, *X*.
b. there is a large variance of the error term, u.
c. the sample size is smaller.
d. the intercept, β_0, is small.

5. Among all unbiased estimators that are weighted averages of $Y_1, \ldots, Y_n, \overline{Y}$ is

a. the only consistent estimator of μ_Y.
b. the most efficient estimator of μ_Y.
c. a number which, by definition, cannot have a variance.
d. the most unbiased estimator of μ_Y.

6. Consider the following regression line: $\widehat{TestScore} = 698.9 - 2.28 \times STR$. You are told that the *t*-statistic on the slope coefficient is 4.38. What is the standard error of the slope coefficient?

a. 0.52
b. 1.96
c. -1.96
d. 4.38

Part B – Short Answer – 6 marks each

7. Derive the variance of the sample average, $\bar{Y}$. Describe two different reasons why we would want to know this variance.

8. Explain what L.S.A. #1 means, and why it is important for OLS estimation.

9. Suppose that you have data on the hourly wages of 100 male workers, and 100 female worker. The sample average of the wages for males is \$16.2, and the sample average of the wages for females is \$14.5. Using a dummy variable (binary variable), write down a population model that you could use to estimate the wage gender gap. What are the OLS estimates for your population model?

10. Describe how the central limit theorem is useful in determining the sampling distribution of the OLS estimator. (Hint: You may want to look at the formula for $\hat{\beta}_1$.)

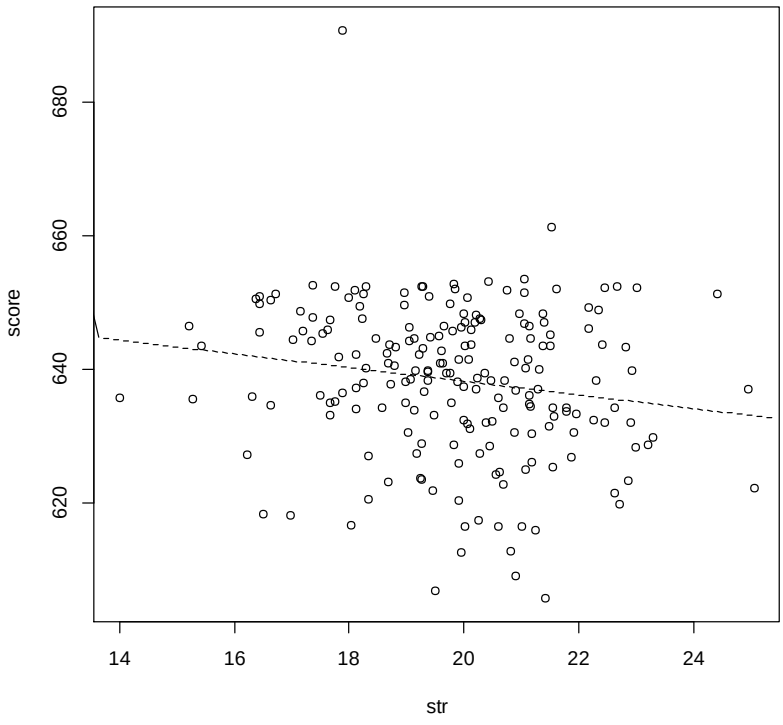
Part C – Long Answer – each part is worth 5 marks

11. The following question uses a subset (*n* = 200) of the California test score data.

The fitted model, estimated by OLS, is:

$$\widehat{score}_i = \underset{(8.41)}{658.65} - \underset{(0.42)}{1.06} \times str_i \quad , \quad R^2 = 0.029$$

Below is a plot of the data, where the dashed line indicates the OLS fitted model:



a) Does the value for R^2 mean that str is not an important variable? Explain.

b) Conduct a formal hypothesis test to determine whether or not str effects $score$.

c) If $s_Y^2 = 128.6$, what is ESS (the explained sum-of-squares)?

d) What does the model predict for $score$, given a class size of 20?

e) Comment on the interpretation of the intercept in this model.

f) Construct a 95% confidence interval around $\hat{\beta}_1$.

END