Department of Economics
University of Maryland

John C. Chao
Spring 2015

Economics 422

Midterm Examination

This exam contains 10 multiple-choice questions and two numerical/mathematical problems. You should try to do all questions and problems. Partial credit will be given for the numerical/mathematical problems, so please show all your work for these problems. **GOOD LUCK!**

Multiple-Choice Questions

1. (6 points) 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. (6 points) The p-value associated with testing the null hypothesis

$$H_0 : \mu_Y = 5$$

against the alternative hypothesis

$$H_0 : \mu_Y \neq 5$$

is defined as follows:

(a) $p = 0.05$

(b) $\Pr_{H_0}\left[\left|\overline{Y} - 5\right| > \left|\overline{Y}^{act} - 5\right|\right].$

(c) $\Pr\left[Z > 1.96\right]$, where $Z \sim N(0,1)$.

(d) $\Pr_{H_0}\left[\left|\overline{Y} - 5\right| < \left|\overline{Y}^{act} - 5\right|\right].$

3. (6 points) An estimator $\widehat{\mu}_Y$ of the population value μ_Y is consistent if

(a) $\widehat{\mu}_Y \xrightarrow{p} \mu_Y$.

(b) its mean square error is the smallest possible.

(c) Y is normally distributed.

(d) $\overline{Y} \xrightarrow{p} 0$.

4. (6 points) The t-statistic for testing the null hypothesis

$$H_0 : \mu_Y = 1$$

against the alternative hypothesis

$$H_0 : \mu_Y \neq 1$$

is defined as follows:

(a) $t = \frac{\overline{Y}-1}{\sigma_Y^2/n}$.

(b) $t = \frac{\overline{Y}-1}{SE(\overline{Y})}$.

(c) $t = \frac{\left(\overline{Y}-1\right)^2}{SE(\overline{Y})}$

(d) 1.96.

5. (6 points) A large p-value implies

(a) rejection of the null hypothesis.

(b) a large t-statistic.

(c) a large $\overline{Y}^{act}$.

(d) that the observed value $\overline{Y}^{act}$ is consistent with the null hypothesis.

6. (6 points) The standard error of the regression (SER) for the simple linear model

$$Y_i = \beta_0 + \beta_1 X_i + u_i,\ i = 1, ..., n,$$

is defined as follows:

(a) $\sqrt{\frac{1}{n-2}\sum_{i=1}^{n} \widehat{u}_i^2}$

(b) SSR

(c) $1 - R^2$

(d) $\sqrt{\frac{1}{n-1}\sum_{i=1}^{n} \widehat{u}_i^2}$

Here, $\widehat{u}_i$ $(i = 1, ..., n)$ denote the OLS residuals.

7. (6 points) The following are all least squares assumptions with the exception of

(a) The conditional distribution of u_i given X_i has a mean of zero with probability one.

(b) The explanatory variable in the regression model is normally distributed.

(c) (X_i, Y_i), $i = 1, ..., n$ are independent and identically distributed.

(d) Large outliers are unlikely.

8. (6 points) Which of the following statements is correct?

(a) $TSS = ESS + SSR$

(b) $ESS = SSR + TSS$

(c) $ESS > TSS$

(d) $R^2 = 1 - (ESS/TSS)$

9. (6 points) The OLS estimator is derived by

(a) connecting the Y_i corresponding to the lowest X_i observation with the Y_i corresponding to the highest X_i observation.

(b) making sure that the standard error of the regression equals the standard error of the slope estimator.

(c) minimizing the sum of absolute residuals.

(d) minimizing the sum squared residuals.

10. (6 points) The normal approximation to the sampling distribution of $\widehat{\beta}_1$ is powerful because

(a) many explanatory variables in real life are normally distributed.

(b) it allows econometricians to develop methods for statistical inference without explicitly knowing the underlying distribution of the data.

(c) many other distributions are not symmetric.

(d) implies that OLS has small variance.

Problems:

11. (20 points)

Suppose you are interested in how well ones height predicts ones weight. Hence, you collect data on 110 people in order to run the simple regression

$$Y_i = \beta_0 + \beta_1 X_i + u_i,\ i = 1, ..., 110;$$

where Y_i denotes the weight of the i^{th} individual and X_i denotes the height of the i^{th} individual. From this data, you calculated the following quantities:

$$\sum_{i=1}^{110} Y_i = 17,375,\ \sum_{i=1}^{110} \left(Y_i - \overline{Y}\right)^2 = 94,228.8,$$

$$\sum_{i=1}^{110} X_i = 7665.5, \sum_{i=1}^{110} \left(X_i - \overline{X}\right)^2 = 1248.9,$$

$$\sum_{i=1}^{110} \left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right) = 7625.9$$

where $\overline{Y}$ and $\overline{X}$ denote the respective sample means.

(a) From the information given above, calculate the OLS estimates for β_0 and β_1. Show your work.

(b) Calculate the (unadjusted) R^2 measure for this regression and explain its meaning. Show your work.

12. (20 points)

Suppose you have the random sample

$$X_1, X_2, ..., X_n \sim i.i.d.\,(\mu_X, 4)$$

Consider the statistic

$$S_n = \sqrt{n} X_n,$$

where

$$\overline{X}_n = \frac{1}{n} \sum_{i=1} X_i.$$

(a) What is the approximate distribution of S_n under the null hypothesis that $\mu_X = 0$, when the sample size n is large? Explain how you have arrived at your answer.

(b) Calculate an approximate value for the probability

$$\Pr\left[-1.25 \leq \overline{X}_n - 1 \leq -0.5\right]$$

under the null hypothesis that $\mu_X = 0$, when $n = 100$. Show your work.