EC 122: Econometrics
California Institute of Technology
Winter 2006

Professor: Tae-Hwy Lee (Email: tlee@hss.caltech.edu; Phone 626-395-3531; Office: 104 Baxter)
Lecture: MW 10:30-11:55 a.m., 127 Baxter
Office Hours: M 2:00-3:30 p.m. or by appointment

TA: Jihong Lee (Email: jhl@hss.caltech.edu; Office: 6 Baxter)
Office Hours: MW 3:00-4:00 p.m. or by appointment

Course Description: The course is an introduction to econometrics. The course covers the basic methods and theory of modern econometrics while it is also intended to provide a foundation for applied research in economics. Numerous empirical examples together with rigorous derivations will be discussed. Familiarity with probability, statistics, and matrix algebra is assumed. The objective of the course is to learn essential statistical methods to analyze the relationship between two or more variables. It is designed to be useful for data analysis in economics. High degree of student participation is expected in the form of questions and answers relating to the material being developed in lectures. Class attendance is required.

Exams and Grading: There will be weekly homework assignments (50%) and a final exam (50%). The weekly homework assignments will be considered as multiple take-home midterm exams. The final exam is comprehensive. Late homework will not be considered. The assignments and the final exam will be graded by the TA and will then be reviewed by the instructor. You are fully responsible for following up all the announcements made during the lectures.

Textbook: *Required.* Wooldridge, Jeffrey M., 2006, *Introductory Econometrics: A Modern Approach, 3*ed., South-Western Publisher of Thomson. You are expected to have the textbook in the classroom throughout the entire term as we discuss numerous examples from the textbook. The textbook does not provide all the proofs of many important theoretical results, which will be provided in the classroom lecture.

Course Outline: We will follow the text fairly closely (but occasionally the coverage of the material will be different from that in the text).

Regression Analysis with Cross-Sectional Data	
Chapter 2	Simple Linear Regression Model
Chapter 3	Multiple Linear Regression: Estimation
Appendix E	Linear Regresion with Matrix Algebra
Chapter 4	Multiple Linear Regression: Inference
Chapter 5	Multiple Linear Regression: Asymptotics
Chapter 6	Multiple Linear Regression: Further Issues
Chapter 7	Multiple Linear Regression: Binary Variables
Chapter 8	Heteroskedasticity
Chapter 9	More on Specification and Data Problems
Regression Analysis with Time Series Data	
Chapter 10	Regression with Stationary Time Series
Chapter 11	Regression with Non-stationary Time Series
Chapter 12	Time Series Regression with Serial Correlation and Heteroskedasticity
Chapter 18	Advanced Time Series Topics

Some rules: You are fully responsible for following up all the announcements made during the lectures. You are required to attend all lectures. You are required to read the textbook. All due dates are firm. No make-up exams. Come to the class before the class starts. Each class will start on time. Please do not leave the room after the lecture has started. If possible, please do not email to the instructor and the TA. Use their office hours and talk to them in person. Turn off cell phones during the lecture.

EC 122: HW1

1. (20) Consider the simple linear regression (SLR) model $y_i = \beta_0 + \beta_1 x_i + u_i$. Is each of the following statement true or false?

   (a) $\sum_{i=1}^n u_i = 0$?
   (b) $\sum_{i=1}^n \hat{u}_i = 0$?
   (c) $\sum_{i=1}^n x_i u_i = 0$?
   (d) $\sum_{i=1}^n x_i \hat{u}_i = 0$?
   (e) $\sum_{i=1}^n y_i u_i = 0$?
   (f) $\sum_{i=1}^n \hat{y}_i u_i = 0$?
   (g) $\sum_{i=1}^n y_i \hat{u}_i = 0$?
   (h) $\sum_{i=1}^n \hat{y}_i \hat{u}_i = 0$?
   (i) $\sum_{i=1}^n \bar{y} \hat{u}_i = 0$?
   (j) $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$?
   (k) $\bar{y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{x}$?
   (l) $\bar{y} = \beta_0 + \beta_1 \bar{x}$?
   (m) $y_i = \hat{y}_i + \hat{u}_i$ for all i?
   (n) $\sum_{i=1}^n y_i = \sum_{i=1}^n \hat{y}_i + \sum_{i=1}^n \hat{u}_i$?
   (o) $\sum_{i=1}^n y_i^2 = \sum_{i=1}^n \hat{y}_i^2 + \sum_{i=1}^n \hat{u}_i^2$?
   (p) $\sum_{i=1}^n (y_i - \bar{y}) = \sum_{i=1}^n (\hat{y}_i - \bar{y}) + \sum_{i=1}^n \hat{u}_i$?
   (q) $\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n \hat{u}_i^2$?
   (r) $\sum_{i=1}^n (y_i - \bar{y}) = 0$?
   (s) $\sum_{i=1}^n (\hat{y}_i - \bar{y}) = 0$?
   (t) $R^2 = 1 - SSR/SST$?

2. (20) Consider the SLR model $y_i = \beta_0 + \beta_1 x_i + u_i$.

   (a) Derive the OLS estimators $\hat{\beta}_j$ $(j = 0, 1)$.
   (b) Explain why $\hat{\beta}_1$ is a random variable.
   (c) What is the mean of $\hat{\beta}_1$?
   (d) What is the variance of $\hat{\beta}_1$?
   (e) Prove that $\hat{\beta}_1$ is an unbiased estimator. What is the bias of $\hat{\beta}_1$?
   (f) It is known that $\hat{\beta}_1$ is the BLUE. What does this mean?
   (g) Define SSR, SST, and SSE. Show that $SST = SSE + SSR$.

3. (20) The following questions are regarding the Assumptions for SLR model. State each of the assumptions SLR 1-5, and then answer each of the following questions.

   (a) Why is Assumption SLR.1 important? Give a situation when this assumption may fail.
   (b) Why is Assumption SLR.2 important? Give a situation when this assumption may fail.
   (c) Why is Assumption SLR.3 important? Give a situation when this assumption may fail.
   (d) Why is Assumption SLR.4 important? Give a situation when this assumption may fail.
   (e) Why is Assumption SLR.5 important? Give a situation when this assumption may fail.

4. (20) The following questions are regarding the assumptions for SLR model.

(a) Under which of Assumptions SLR.1-5, $E(\hat{\beta}_1) = \beta_1$?

(b) Under which of Assumptions SLR.1-5, $Var(\hat{\beta}_1) = \sigma^2/SST_x$?

(c) Using the law of iterated expectations, show that Assumption SLR.3 $E(u|x) = 0$ implies $E(u) = 0$.

(d) Using the law of iterated expectations, show that Assumption SLR.3 $E(u|x) = 0$ implies $E(xu) = 0$.

(e) Explain how you use the above two results in (c) and (d) for the method of moment to estimate the two unknown parameters β_0 and β_1.

5. (20) Answer the following questions, using the Eviews output below, for the SLR model

$$PRATE_i = \beta_0 + \beta_1 MRATE_i + u_i.$$

(a) What is R^2?

(b) What is SSR?

(c) What is SST?

(d) What is SSE?

(e) What is $\hat{\beta}_0$?

(f) What is $\hat{\beta}_1$?

(g) What is $\bar{y}$?

(h) What is $Var(\hat{\beta}_0)$?

(i) What is $Var(\hat{\beta}_1)$?

(j) What is n?

Dependent Variable: PRATE
Method: Least Squares
Sample: 1 1534
Included observations: 1534

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	83.07546	0.563284	147.4840	0.0000
MRATE	5.861079	0.527011	11.12137	0.0000

R-squared	0.074703	Mean dependent var	87.36291
Adjusted R-squared	0.074099	S.D. dependent var	16.71654
S.E. of regression	16.08528	Akaike info criterion	8.394989
Sum squared resid	396383.8	Schwarz criterion	8.401945
Log likelihood	-6436.956	F-statistic	123.6848
Durbin-Watson stat	1.908008	Prob(F-statistic)	0.000000

EC 122: HW2

Distributed on January 23 (M), Due on February 1 (W) at 10:29 a.m.

1. (20) We discussed a simple (location) model $y_i = \beta_0 + u_i$, in the classes on January 11 and 18. Prove Properties 1, 2, 3, 4, 5, for this model, as stated on January 18 in the class. Make necessary assumptions *where* you need in proving for each of the five properties.
   [Note: On January 18 we did not prove these properties. Instead, yesterday on January 22 (Sunday), we proved five theorems (each corresponding to the five properties on January 18) for SLR model $y_i = \beta_0 + \beta_1 x_{i1} + u_i$.]

2. (20) Consider the location model $y_i = \beta_0 + u_i$. Is each of the following estimators for β_0 unbiased? State *explicitly* assumptions you make to arrive at your conclusions.

   (a) $\hat{\beta}_0 = y_1$
   (b) $\hat{\beta}_0 = \frac{1}{2}(y_1 + y_n)$
   (c) $\hat{\beta}_0 = n^{-1}\sum_{i=1}^n y_i$
   (d) $\hat{\beta}_0 = \left(\frac{1-a}{1-a^n}\right)\sum_{i=1}^n a^{i-1} y_i$ where $0 < a < 1$
   (e) $\hat{\beta}_0 = \bar{y} + \frac{1}{n}$ where $\bar{y} = n^{-1}\sum_{i=1}^n y_i$

3. (20) Consider the SLR model $y_i = \beta_0 + \beta_1 x_{i1} + u_i$. For each of the following testing hypotheses, draw the power function of the t test and show that the test is "consistent".

   (a) $H_0 : \beta_1 = c$ vs $H_1 : \beta_1 \neq c$ (where c is a given constant).
   (b) $H_0 : \beta_1 = c$ vs $H_1 : \beta_1 > c$ (where c is a given constant).
   (c) $H_0 : \beta_1 = c$ vs $H_1 : \beta_1 < c$ (where c is a given constant).

4. (20) Consider the MLR model $y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + u_i$.

   (a) For $k = 2$, derive Equation (3.22). See page 119.
   (b) For $k = 2$, derive Equation (3.23). See pages 120-121
   (c) State the GAUSS-Markov Theorem with necessary assumptions.
   (d) How would you estimate the variance $\sigma^2 = var(u|x_1, \ldots, x_k)$?
   (e) What is the meaning of "perfect multicollinearity"?
   (f) How does each of the following affect $var(\hat{\beta}_1|\mathbf{X})$ (where $\mathbf{X}$ is defined on January 22 in the class)?
      i. $var(u|x_1, \ldots, x_k)$.
      ii. $corr(x_1, x_2)$.
      iii. $var(x_1)$.
      iv. $var(x_2)$.

5. (20) Explain the consequences of the following in terms of unbiasedness and efficiency of the OLS estimators.

   (a) Omitting an important variable.
   (b) Including an irrelevant variable.

6. (20) Consider the MLR model $y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + u$.

   (a) State the MLR assumptions (MLR.1, MLR.2, MLR.3, MLR.4, MLR.5, MLR.6).
   (b) Suppose that u is independent of the explanatory variables, and it takes the values $-2, -1, 0, 1,$ and 2 with equal probability of 0.2. Does this violate each of the MLR assumptions? Explain.

7. (20) Answer the following questions, using the Eviews output below, for the model

$$colGPA_i = \beta_0 + \beta_1 hsGPA_i + \beta_2 ACT_i + u_i.$$

(a) What are SSR, SST, SSE of this regression?

(b) What is n?

(c) What is k?

(d) What is $\frac{1}{n}\sum_{i=1}^{n} colGPA_i$?

(e) What is the values of $se(\hat{\beta}_1)$?

(f) Test for $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 > 0$. What is the statistic? What is the degrees of freedom of your statistic? What is the p-value? What is your conclusion at the 5% level?

(g) Test for $H_0 : \beta_2 = 0$ vs $H_1 : \beta_2 \neq 0$. What is the statistic? What is the degrees of freedom of your statistic? What is the p-value? What is your conclusion at the 5% level?

Dependent Variable: COLGPA
Method: Least Squares
Sample: 1 141
Included observations: 141

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.286328	0.340822	3.774191	0.0002
HSGPA	0.453456	?	4.732721	0.0000
ACT	0.009426	0.010777	0.874627	0.3833

R-squared	0.176422	Mean dependent var	3.056738
Adjusted R-squared	0.164486	S.D. dependent var	0.372310
S.E. of regression	0.340316	Akaike info criterion	0.703161
Sum squared resid	15.98244	Schwarz criterion	0.765901
Log likelihood	-46.57287	F-statistic	14.78073
Durbin-Watson stat	1.885351	Prob(F-statistic)	0.000002

EC 122: HW3

Distributed on January 30 (M). Due on February 8 (W) at 10:29 a.m.

1. (20) Consider the model $Y_i = \beta_0 + u_i,\ i = 1, \cdots, n,\ u_i \sim i.i.d.\ N(0, \sigma^2)$. Let $\hat{\beta}_0 = n^{-1} \sum y_i \equiv \bar{y}$, $\hat{\sigma}^2 = n^{-1} \sum (y_i - \bar{y})^2$ and $s^2 = (n-1)^{-1} \sum (y_i - \bar{y})^2$.
   (a) Show that $\hat{\beta}$ is the maximum likelihood estimator. [Also check the second order condition.]
   (b) Show that $\text{var}(\hat{\beta}_0)$ attains the CRLB.
   (c) Show that $\hat{\sigma}^2$ is biased and s^2 is unbiased.
   (d) Show that $n\hat{\sigma}^2/\sigma^2$ and $(n-1)s^2/\sigma^2$ follows a chi-squared distribution with $(n-1)$ degrees of freedom.

2. (20) Production data for 22 firms in a certain industry produce the following, where $y = \ln(\text{output})$ and $x = \ln(\text{labor hours input})$: $\bar{y} = 20$, $\bar{x} = 10$, $\sum_{i=1}^n (y_i - \bar{y})^2 = 100$, $\sum_{i=1}^n (x_i - \bar{x})^2 = 60$, $\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = 30$.
   (a) Compute the ordinary LS (OLS) estimates of β_0 and β_1 in the model $y = \beta_0 + \beta_1 x + u$.
   (b) Test the hypothesis that $\beta_1 = 1$.
   (c) Form a 95% confidence interval for σ^2, the variance of u.

3. (20) Consider the model $y_i = \beta_0 + \beta_1 x_i + u_i$. Explain whether and why the following is true or false.
   (a) If $E(u_i|x_i) = 0$, then $E(x_i u_i) = 0$.
   (b) If $E(u_i|x_i) = 0$, then $E(x_i^2 u_i) = 0$.
   (c) If $E(u_i|x_i) = 0$, then u_i is independent of x_i.

4. (20) Consider the model

$$\log(wage) = \beta_0 + \beta_1 jc + \beta_2 univ + \beta_3 exper + u,$$

   where *jc* is number of years attending a two-year college and *univ* is number of years at a four-year college. Suppose we want to test a hypothesis whether one year at a junior college is worth one year at a university. State the null and alternative hypotheses. Explain how you would conduct the test.

5. (20) Consider the model $Y = X\beta + \varepsilon$ and the OLS estimator $\hat{\beta}_n = (X'X)^{-1}X'Y$.
   (a) Under what minimal assumptions is $\hat{\beta}$ unbiased? Prove it under the minimal assumptions.
   (b) Under what minimal assumptions is $\hat{\beta}$ the best linear unbiased estimator? Prove it under the minimal assumptions.

6. (20) Prove that the OLS estimators of β in the following linear models are identical

$$y_t = a + bt + \beta x_t + \varepsilon_t \qquad t = 1, \ldots, n,$$

$$y_t^* = \beta x_t^* + \varepsilon_t^*,$$

   where y_t^* and x_t^* are de-trended y_t and x_t, obtained by regressing y_t and x_t on t and setting y_t^* and x_t^* equal to the respective residuals. Also, show whether ε_t^* and ε_t are the same or not.

7. (20) Consider a stochastic process $Z_t \equiv (Y_t\ X_t')'$ where Y_t is the variable of interest and X_t is a vector of other variables. Suppose that there is a decision maker who takes an one-step point forecast $f_{t,1} \equiv f(Z_t, \beta)$ of Y_{t+1} and uses it in some relevant decision. The one-step forecast error $e_{t+1} \equiv Y_{t+1} - f_{t,1}$ will result in a cost of $c(e_{t+1})$, where the function $c(e)$ will increase as e increases in size, but not

necessarily symmetrically or continuously. The optimal forecast $f^*_{t,1}$ will be chosen to produce the forecast errors that minimize the expected loss

$$\min_{f_{t,1}} \int_{-\infty}^{\infty} c(y - f_{t,1}) dF_t(y),$$

where $F_t(y) \equiv \Pr(Y_{t+1} \leq y | I_t)$ where I_t is some proper information set at time t, and so includes Z_{t-j}, $j \geq 0$. The corresponding optimal forecast error will be

$$e^*_{t+1} = Y_{t+1} - f^*_{t,1.}$$

Consider predicting Y_t given X_t.

(a) Let $c(e_{t+1}) = e^2_{t+1}$. Show that the conditional expectation of Y_{t+1} given I_t, denoted as $E(Y_{t+1}|I_t)$, gives the minimum mean squared error prediction of Y_{t+1} based on I_t.

(b) Show that $\varepsilon_{t+1} = Y_{t+1} - E(Y_{t+1}|I_t)$ forms a martingale difference sequence with respect to the information set I_t.

(c) Discuss how the above two results would be affected if the objective function $c(\cdot)$ is the check function.

8. (20) Under the classical assumptions on the model $Y = X\beta + u$, where β is to be estimated using the squared-error loss $c(u_i) = u_i^2$.

(a) Show that R^2 is the squared correlation between Y and $\hat{Y} = X\hat{\beta}_n$.

(b) Show that $E(\hat{u}'\hat{u}) = (n - k - 1)\sigma^2$, where $\hat{u} = Y - \hat{Y}$.

(c) Show that $\hat{\beta}$ and $\hat{u}$ are independent.

EC 122: HW4

Distributed on February 6 (M). Due by February 15 (W) at 10:29 a.m.

1. (20) Prove that the unadjusted coefficient of determination R^2 is non-decreasing when additional regressors are added.

2. (20) Consider the model $Y = X\beta + u$ where $u|X \sim N(\mathbf{0}, \sigma^2 I_n)$. Is each of the following 11 statements true or false? Give full explanation for your answer.

   (a) $E(u'u|X) = \sigma^2, E(u\,u'|X) = \sigma^2, E(\hat{u}'\hat{u}|X) = \sigma^2, E(\hat{u}\hat{u}'|X) = \sigma^2$

   (b) $u'X = 0, \hat{u}'X = 0, \hat{u}'Y = 0, \hat{u}'\hat{Y} = 0, u'\hat{Y} = 0$

   (c) $\hat{\beta}$ and $\hat{\sigma}^2$ are independent.

   (d) $\hat{\beta}$ and $\hat{\varepsilon}$ are independent.

3. (20) Consider $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + u_i, i = 1, \cdots, n = 100$.

$$\sum (Y_i - \bar{Y})^2 = 493/3,\ \sum (X_{i1} - \bar{X}_1)^2 = 30,$$

$$\sum (X_{i2} - \bar{X}_2)^2 = 3,\ \sum (X_{i1} - \bar{X}_1)(Y_i - \bar{Y}) = 30,$$

$$\sum (X_{i2} - \bar{X}_2)(Y_i - \bar{Y}) = 20,\ \sum (X_{i1} - \bar{X}_1)(X_{i2} - \bar{X}_2) = 0.$$

   (a) Compute the OLS estimates of β_1, β_2, and R^2.

   (b) Test $H_0 : \beta_1 = \beta_2 = 0$.

   (c) Test $H_0 : \beta_1 = 7\beta_2$.

4. (20) Consider $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + u_i, i = 1, \cdots, n$. Given the following results, compute $n, \bar{X}_1, \bar{X}_2, \hat{\sigma}^2$, and t values for $\beta_j,\ j = 0, 1, 2, 3$.

$$X'X = \begin{bmatrix} 10 & & & \\ 40 & 200 & & \\ 80 & 330 & 710 & \\ 50 & 220 & 390 & 270 \end{bmatrix} \qquad (X'X)^{-1} = \begin{bmatrix} 6.69 & & & \\ 0.51 & 0.09 & & \\ -0.46 & -0.04 & 0.04 & \\ -0.99 & -0.11 & 0.06 & 0.19 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} 100 \\ 419 \\ 792 \\ 506 \end{bmatrix} \qquad (X'X)^{-1}X'Y = \begin{bmatrix} 17.43 \\ 1.37 \\ -0.72 \\ -1.43 \end{bmatrix} \qquad \hat{u}'\hat{u} = 22.79.$$

5. (20) Consider $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + u_i$, where $i = 1, \cdots, n,\ (n = 5)$ and $u_i|X_i \sim N(0, \sigma^2)$. The data are given by

$$Y = \begin{bmatrix} 3 \\ 1 \\ 8 \\ 3 \\ 5 \end{bmatrix} \quad \text{and} \quad X = \begin{bmatrix} 1 & 3 & 5 \\ 1 & 1 & 4 \\ 1 & 5 & 6 \\ 1 & 2 & 4 \\ 1 & 4 & 6 \end{bmatrix}.$$

   (a) Test $H_0 : \beta_1 = \beta_2 = 0$.

   (b) Test $H_0 : \beta_1 + \beta_2 = 0$.

6. (20) Consider the following production function model

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + u_i, i = 1, \cdots, n.$$

where Y is log output, X_1 is log labor input, X_2 is log capital input, and the subscript i refers to the ith firm. All assumptions of the classical linear regression model are supposed to be satisfied. The information from a random sample of 23 firms is summarized as

$$\sum(Y_i - \bar{Y})^2 = 10, \ \sum(X_{i1} - \bar{X}_1)^2 = 12,$$

$$\sum(X_{i2} - \bar{X}_2)^2 = 12, \ \sum(X_{i1} - \bar{X}_1)(Y_i - \bar{Y}) = 10,$$

$$\sum(X_{i2} - \bar{X}_2)(Y_i - \bar{Y}) = 8, \text{ and } \sum(X_{i1} - \bar{X}_1)(X_{i2} - \bar{X}_2) = 8.$$

(a) Find the least squares estimates of β_1 and β_2. Calculate the value of R^2 and the sum of squares of residuals.

(b) You are told that $\beta_1 + \beta_2 = 1$, i.e., constant returns to scale. You wish to assume constant returns a priori and specify

$$\beta_2 = 1 - \beta_1.$$

Obtain the least squares estimate of β_1 under the restriction. Calculate the value of R^2 and the sum of squares of residuals.

(c) Using the above results, carry out a test for constant returns to scale, i.e., test the hypothesis $H_0 : \beta_1 + \beta_2 = 1$ against the two-sided alternative, using an F test. Discuss what you find.

7. (20) Show the following:

(a) Show that the sample mean of a variable may be obtained by running a regression of the variable on a constant and by reading the estimate of the constant term.

(b) Show that the difference between the sample means of the two groups may be obtained by running a regression of the variable on a constant and a dummy variable (a binary variable) that indicates whether an observation is from a group and by reading the coefficient on the dummy variable.

EC 122: HW5

Distributed on February 13 (M). Due **by February 27 (M)** at 10:29 a.m.

1. (40) Consider the location model $y_i = \beta_0 + u_i$. Is each of the following estimators for β_0 consistent? State *explicitly* assumptions you make to arrive at your conclusions.

   (a) $\hat{\beta}_0 = y_1$: [Answer: This is unbiased but not consistent as $\Pr(|\hat{\beta}_0 - \beta_0| \geq \delta) = \Pr(|y_1 - \beta_0| \geq \delta)$, which must be positive for sufficiently small δ (because $\sigma^2 > 0$), so $\Pr(|\hat{\beta}_0 - \beta_0| \geq \delta)$ does not tend to zero as $n \to \infty$, so $\hat{\beta}_0 = y_1$ is not consistent.]

   (b) $\hat{\beta}_0 = \frac{1}{2}(y_1 + y_n)$

   (c) $\hat{\beta}_0 = n^{-1} \sum_{i=1}^{n} y_i$

   (d) $\hat{\beta}_0 = \left(\frac{1-a}{1-a^n}\right) \sum_{i=1}^{n} a^{i-1} y_i$ where $0 < a < 1$. [Hint: This is unbiased but not consistent. Show why.]

   (e) $\hat{\beta}_0 = \bar{y} + \frac{1}{n}$ where $\bar{y} = n^{-1} \sum_{i=1}^{n} y_i$. [Hint: This is biased but consistent. Show why.]

2. (60) Derive the Wald and Lagrange multiplier (LM) statistics. Find their limiting distributions under the null hypotheses.

   (a) $Y_i = \beta_0 + u_i$. Test $H_0 : \beta_0 = 1$.

   (b) $Y_i = \beta_0 + \beta_1 X_{i1} + u_i$. Test $H_0 : \beta_1 = 0$.

   (c) $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + u_i$. Test $H_0 : \beta_1 = \beta_2 = 0$.

   (d) $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + u_i$. Test $H_0 : \beta_1 = \beta_2$.

   (e) $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + u_i$. Test $H_0 : \beta_1 + \beta_2 = 1$.

   (f) $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + u_i$. Test $H_0 : \beta_1 \beta_2 = 1$.

The Wald test is not invariant to the way the hypothesis is formulated. Think about how this can happen.

Suppose you want to estimate the following dynamic model using the U.S. macroeconomic data

$$\ln C_t = \alpha + \beta \ln Y_t + \gamma \ln C_{t-1} + u_t$$

where $u_t | Y_t, C_{t-1} \sim (0, \sigma^2)$. Let $\theta = (\alpha\ \beta\ \gamma)'$ and its consistent estimator $\hat{\theta}_n = (\hat{\alpha}_n\ \hat{\beta}_n\ \hat{\gamma}_n)'$.

1. Let $c(\theta) = \frac{\beta}{1-\gamma}$ be the long run marginal propensity to consume. Consider testing the hypothesis $H_0 : c(\theta) = 1$. Compute
$$Wald_1 = n(c(\hat{\theta}_n) - 1)'\hat{\Gamma}_n^{-1}(c(\hat{\theta}_n) - 1)$$
where $\hat{\Gamma}_n = \nabla c(\hat{\theta}_n)'\hat{D}_n \nabla c(\hat{\theta}_n)$, $\hat{D}_n = \hat{\sigma}_n^2 \hat{M}_n$, $\hat{M}_n = X'X/n$, $\hat{\sigma}_n = \hat{u}'\hat{u}/n$, and $\nabla c(\theta) = \frac{\partial c(\theta)}{\partial \theta'}$.

2. $H_0 : c(\theta) = \frac{\beta}{1-\gamma} = 1$ implies $H_0' : \beta + \gamma = 1$. Note that while H_0 is a nonlinear restriction, H_0' is a linear restriction which is of the form $R\theta = r$ with $R = (0\ 1\ 1)$ and $r = 1$. Compute the Wald test statistic for H_0'
$$Wald_2 = n(R\hat{\theta}_n - 1)'\hat{\Gamma}_n^{-1}(R\hat{\theta}_n - 1)$$
where $\hat{\Gamma}_n = R\hat{D}_n R'$, $\hat{D}_n = \hat{\sigma}_n^2 \hat{M}_n$, and $\hat{M}_n = X'X/n$.

3. In general $Wald_1$ and $Wald_2$ may not be the same.

Testing a hypothesis when Assumption (6) fails or when Assumption (6') is not appropriate as n is not large enough: Bootstrap

Car dealers across North America use theBluebook to help them determine the value of used cars that their customers trade in when purchasing new cars. The book, which is published monthly, list the trade-in values for all basic models of cars. It provides alternative values of each car model according to its condition and optional features. The values are determined on the basis of the average paid at recent used-car auctions. (These auctions are the source of supply for many used-car dealers.) However, the Bluebook does not indicate the value determined by the odometer reading, despite the fact that a critical factor for used-car buyers is how far the car has been driven. To examine this issue, a used-car dealer randomly selected 30 three-year old Ford Tauruses that were sold at auction during the past month. Each car was in top condition and equipped with automatic transmission, AM/FM cassette player, and air conditioning. The dealer recorded the price and the number of miles on the odometer. The dealer wants to find the regression line, $y_i = a + bx_i + u_i$, i.e.,

$$(\text{Auction Selling Price})_i = a + b(\text{Odometer Reading})_i + u_i, \quad i = 1, \ldots, n = 30.$$

We estimate the model by OLS. We want to examine from these data that higher mileages result in lower price. Suppose that Assumption (5) holds. Let F_0 be the conditional density of the error term u_i given the regressor x_i. Suppose we want to test for $H_0 : b = 0$ versus $H_1 : b < 0$ using the t statistic, denoted as $T_n = \hat{b}_n / se(\hat{b}_n)$. We need to get the distribution of T_n, denoted as $G_n(\tau, F_0) \equiv \Pr(T_n < \tau | F_0)$.

1. $G_n(\tau, F_0)$: If we may make Assumption (6) that F_0 is normal, T_n follows the Student-t distribution with $n-2$ degrees of freedom, i.e., $T_n \sim t(n-2)$. The Student-t distribution is to take $G_n(\tau, F_0)$ with F_0 being assumed to be normal.

2. $G_\infty(\tau, F_0)$: However, we never know F_0 and thus assuming (6) may be very misleading. Hence, we do not wish to assume (6). When $n \to \infty$, we may use Assumption (6') under which we have the large sample result that $T_n \to^d N(0,1)$ as $n \to \infty$.

3. $G_n(\tau, F_n)$: When neither Assumption (6) or (6') is good, we can use the bootstrap method. Explain what the bootstrap is, how you may implement it, and why it would work.

EC 122: HW6

Distributed on February 22 (W). Due by March 1 (W), 10:29 a.m.

HW6 is about Assumption (1). Read Chapters 4 and 5. Suppose Assumption (6) may not hold. Consider $Y = X\beta + u = X_1\beta_1 + X_2\beta_2 + u$, where β is $(k+1)\times 1$, β_1 is $k_1 \times 1$, β_2 is $k_2 \times 1$, $k+1 = k_1 + k_2$, $E(u|X) = 0$, and $E(uu'|X) = \sigma^2 I_n$. Let $\hat{\beta}_1$ and $\hat{\beta}_2$ denote the estimators from the regression $Y = X_1\hat{\beta}_1 + X_2\hat{\beta}_2 + \hat{u}$. Let β_1^* denote the estimator from the regression $Y = X_1\beta_1^* + u^*$.

1. (10) Suppose $\beta_2 \neq 0$. Show that $\hat{\beta}_1$ is unbiased for β_1, while β_1^* is biased unless $E(X_1'X_2) = 0$.
2. (10) Suppose $\beta_2 = 0$. Show that $\hat{\beta}_1$ and β_1^* are both unbiased for β_1, while β_1^* is more efficient. [Hint: We have done these #1 and #2 several weeks ago. For #2 on efficiency, instead of deriving the variances, simply use the Gauss-Markov theorem.]
3. (20) Hence we want to test $H_0 : \beta_2 = 0$. Show that the LM statistic for $H_0 : \beta_2 = 0$ is nR_u^2, which is distributed asymptotically as $\chi^2_{k_2}$ under H_0 when Assumption (5) holds, where R_u^2 is obtained from regressing u^* on X (that is X_1 and X_2). Discuss why this simple form of LM statistic nR_u^2 is invalid when Assumption (5) fails. [Hint: This is based on the lecture and lecture note on February 15.]
4. (60) Consider a problem of model selection, where you want to compare two nested models,

$$\begin{aligned} Y &= X_1\hat{\beta}_1 + X_2\hat{\beta}_2 + \hat{u}. \\ Y &= X_1\beta_1^* + u^*. \end{aligned}$$

   that is, to consider the question of the inclusion of X_2 in the linear regression. Show whether each of the following **nine** statistics is a "consistent" model selection criterion:

   (a) Test statistics (such as $Wald_n$, LM_n, LR_n).
   (b) Estimation objective functions to minimize (such as SSR) or to maximize (such as $\ln L$ and R^2).
   (c) Penalized objective functions (such as $\bar{R}^2$, AIC, and SIC).

Remarks: We showed (on February 15) that $Wald_n$, LM_n, LR_n are "consistent test statitics" for testing a hypothesis, e.g., $H_0 : \beta_2 = 0$ as in #3, in that their asymptotic power is unity. See my lecture note for Ch 5, page 17. We also showed that $Wald_n$, LM_n, LR_n are asymptotically equivalent in that $LM_n - Wald_n \to^p 0$ (see page 12) and $LR_n - Wald_n \to^p 0$ (see page 16).
In #4, we now show that $Wald_n$, LM_n, LR_n are *not* "consistent model selection criteria" in that they let you select a wrong model with positive probability.

EC 122: HW7

Distributed on March 1 (W). Due by March 8 (W), 10:29 a.m.

1. (30) *This question is about Assumption (1). Read Section 6.2 and Section 9.1 on testing for negelected nonlinearity in the conditional mean:* Consider the model $y_i = X_i'\beta + u_i$, where the parameter vector β is to be estimated under the squared error loss. Suppose you want to check if Assumption (1) holds. Derive a test statistic for the null hypothesis $H_0 : \Pr[E(y_i|X_i) = X_i'\beta] = 1$ for some β, against the alternative hypothesis $H_1 : \Pr[E(y_i|X_i) \neq X_i'\beta] > 0$ for all β. Show that it can be written in the nR_u^2 form under Assumption (5).

2. (20) Problem 9.1 (p. 334). [Read Example 9.1 (p. 306) and Example 9.2 (p. 308) before answering the Problem.]

3. (20) *This question is about Assumption (5). Read Chapter 8 on the White (1980) test for heteroskedasticity:* Consider the model $y_i = X_i'\beta + u_i$, where the parameter vector β is to be estimated under the squared error loss. Suppose you want to check if Assumption (5) holds. Suppose that Assumption (1) holds, i.e., $\Pr[E(y_i|X_i) = X_i'\beta] = 1$ for some β. Show how you would test the null hypothesis that Assumption (5) holds, i.e., $H_0 : \Pr[E(u_i^2|X_i) = \sigma^2] = 1$.

4. (15) Problem 8.1 (p. 298).

5. (15) Problem 8.4 (p. 299). [Read Example 8.1 (p. 274) before answering the Problem.]

EC 122: HW8

Distributed on March 6 (M). Due by March 15 (W), 12 noon

HW8 is about Assumptions (2) and (3'). Read Chapters 10, 11: This problem set is to consider the case when Assumptions (2) and (3') are violated, i.e., $b_n = n^{-1/2}\sum_{t=1}^{n} X_t u_t$ *and* $\hat{M}_n = n^{-1}\sum_{t=1}^{n} X_t X_t'$ *are not* $O_p(1)$.

1. (40) Distribution of the OLS estimator with deterministic trend: Suppose $\{y_t\}_{t=1}^{n}$ are generated from

$$y_t = \alpha + \delta t + u_t$$

where $u_t \sim i.i.d.\ N(0,1)$. Run an OLS regression $y_t = \hat{\alpha}_n + \hat{\delta}_n t + \hat{u}_t$. Show that Assumption (3') fails. Prove that $\hat{\alpha}_n$ and $\hat{\delta}_n$ are consistent for α and δ. Compute the t statistics for $H_0 : \alpha = 1$ and for $H_0 : \delta = 1$. Prove that t_α and t_δ are asymptotically standard normal.

2. (30) Distribution of the OLS estimator with AR(1) with $|\rho| < 1$: Suppose $\{y_t\}_{t=1}^{n}$ are generated from

$$y_t = \rho y_{t-1} + u_t$$

where $u_t \sim i.i.d.\ N(0,1)$ and $\rho = 0.3$. Run an OLS regression $y_t = \hat{\rho}_n y_{t-1} + \hat{\varepsilon}_t$. Does Assumption (3') fail? Prove that $\hat{\rho}_n$ is consistent for ρ. Compute the t statistics for $H_0 : \rho = 0.3$. Prove that t_ρ is asymptotically standard normal.

3. (40) Distribution of the OLS estimator with AR(1) with $\rho = 1$: Suppose that $\{y_t\}_{t=1}^{n}$ are generated from

$$y_t = \rho y_{t-1} + u_t$$

where $u_t \sim i.i.d.\ N(0,1)$ and $\rho = 1$. Run an OLS regression $y_t = \hat{\rho}_n y_{t-1} + \hat{\varepsilon}_t$. Does Assumption (3') fail? Prove that $\hat{\rho}_n$ is consistent for ρ. Compute the t statistics for $H_0 : \rho = 1$. Prove that t_ρ is *not* asymptotically standard normal. Show the limit distribution of t_ρ.

Monte Carlo for HW8

In HW8, you work *analytically*. Here, we repeat HW8 *numerically* via Monte Carlo simulation. The attached are the GAUSS code for the Monte Carlo and its output (four histrograms of the t statistics – t_α and t_δ for #1, t_ρ for #2, and t_ρ for #3).

1. *Distribution of the OLS estimator with deterministic trend:* Generate $\{y_t\}_{t=1}^n$ from

$$y_t = \alpha_0 + \delta_0 t + u_t$$

where $u_t \sim IID\ N(0,1), \alpha_0 = 1$, and $\delta_0 = 1$. Run an OLS regression $y_t = \hat{\alpha}_n + \hat{\delta}_n t + \hat{u}_t$. Compute the t statistics for $H_0 : \alpha = \alpha_0$ and for $H_0 : \delta = \delta_0$. Repeat the above 10,000 times and get t_α and t_δ. Draw histograms of t_α and t_δ. Are they similar to $N(0,1)$? Experiment it with $n = 1000$. Observe that t_α and t_δ are asymptotically standard normal.

2. *Distribution of the OLS estimator with AR(1) with* $|\rho_0| < 1$: Generate $\{y_t\}_{t=1}^n$ from

$$y_t = \rho_0 y_{t-1} + u_t$$

where $u_t \sim IID\ N(0,1)$ and $\rho_0 = 0.3$. Run an OLS regression $y_t = \hat{\rho}_n y_{t-1} + \hat{u}_t$. Compute the t statistics for $H_0 : \rho = 0.3$. Repeat the above 10,000 times and get t_ρ. Draw a histogram of t_ρ. Is it similar to $N(0,1)$? Experiment it with $n = 1000$. Observe that t_ρ is asymptotically standard normal.

3. *Distribution of the OLS estimator with AR(1) with* $\rho_0 = 1$: Generate $\{y_t\}_{t=1}^n$ from

$$y_t = \rho_0 y_{t-1} + u_t$$

where $u_t \sim IID\ N(0,1)$ and $\rho_0 = 1$. Run an OLS regression $y_t = \hat{\rho}_n y_{t-1} + \hat{u}_t$. Compute the t statistics for $H_0 : \rho = 1$. Repeat the above 10,000 times and get t_ρ. Draw a histogram of t_ρ. Is it similar to $N(0,1)$? Experiment it with $n = 1000$. Observe that t_ρ is **not** asymptotically standard normal.

Ec 122: Final Examination
California Institute of Technology
Professor Tae-Hwy Lee
Winter 2006

Instructions:

- Open book, open notes, no collaboration. This examination will not be proctored (as it is a timed take-home exam); you are on the honor system not to cheat or to tolerate cheating by anyone else. There will be no communication with the instructor and the TA; if you believe a question contains an error or an ambiguity, say so on your written examination, make an assumption to the alleged error or to resole the ambiguity, and answer the question as well as you can. Partial credit will be assigned. Please show your work.
- You may take this test during any *consecutive* 4 hour period. Indicate the time/date that you start and the time/date when you stop. Do not look at the exam (next two pages) before you are ready to start the exam.
- Due Wednesday, March 15, by 12:00 noon. Please return to the instructor (104 Baxter), TA (6 Baxter), or Patricia Hamad (112 Baxter).

Ec 122: Final Examination
California Institute of Technology
Professor Tae-Hwy Lee
Winter 2006

Answer all of the following five questions. Each question carries the point as indicated within parenthesis.

1. (20 points) Given the following least squares estimates,

$$\begin{aligned} Y_t &= \text{constant } + 0.92X_t + e_{1t} \\ Y_t &= \text{constant } + 0.84Z_t + e_{2t} \\ Z_t &= \text{constant } + 0.78X_t + e_{3t} \\ X_t &= \text{constant } + 0.55Z_t + e_{4t} \end{aligned}$$

calculate the least squares estimates of β_2 and β_3 in

$$Y_t = \beta_1 + \beta_2 X_t + \beta_3 Z_t + \varepsilon_t.$$

2. (20 points) You want to compare two nested models,

$$\begin{aligned} Y &= X_1\hat{\beta}_1 + X_2\hat{\beta}_2 + \hat{u} \\ Y &= X_1\beta_1^* + u^*, \end{aligned}$$

i.e., to consider the question of the inclusion of X_2 in the linear regression.

(a) (5 points) Show how you test $H_0 : \beta_2 = 0$ using the Lagrange multiplier (LM) test under Assumption (5). Show that the LM test is a consistent test but not a consistent model selection criterion.

(b) (5 points) Show how you test $H_0 : \beta_2 = 0$ using the likelihood ratio (LR) test under Assumption (5). Show that the LR test is a consistent test. Show that the LR test is not a consistent model selection criterion.

(c) (10 points) Note that the above LR test is to compare the likelihoods (without penalizing the size of a model). Now we compare the likelihoods with some penalty terms added. Consider two such penalty terms and form AIC and SIC. Show that SIC is a consistent model selection criterion but AIC is not.

3. (20 points) Suppose $\{y_t\}_{t=1}^n$ are generated from $y_t = \alpha + \delta t + u_t$. Run the OLS regression $y_t = \hat{\alpha}_n + \hat{\delta}_n t + \hat{u}_t$.

(a) (10 points) Show that Assumption (3') fails. Show that $\hat{\alpha}_n$ is $\sqrt{n}$-consistent for α. Show that t_α (the t statistic for $H_0 : \alpha = \alpha_0$) is asymptotically standard normal under H_0. Show that $\hat{\delta}_n$ is $n^{\frac{3}{2}}$-consistent for δ. Explain why $\hat{\delta}_n$ is converging to δ in probability at the rate much faster than the usual rate $n^{\frac{1}{2}}$. Show that t_δ (the t statistic for $H_0 : \delta = \delta_0$) is asymptotically standard normal under H_0.

(b) (10 points) Compute a statistic to test for a joint hypothesis $H_0 : \alpha = \alpha_0$ and $\delta = \delta_0$. Derive the asymptotic distribution of the statistic under H_0.

4. (20 points) Consider the model (in Problem 9.1 of Wooldridge p. 334)

$$\begin{aligned} y_i &= X_i'\boldsymbol{\beta} + u_i \\ &= \beta_0 + \beta_1 \ln(sales_i) + \beta_2 \ln(mktval_i) + \beta_3 profmarg_i + \beta_4 ceoten_i + \beta_5 comten_i + u_i, \end{aligned}$$

where $y_i = \ln(salary_i)$, $X_i' = (1 \ \ln(sales_i) \ \ln(mktval_i) \ profmarg_i \ ceoten_i \ comten_i)$ and the parameter vector $\boldsymbol{\beta}$ is to be estimated under the squared error loss. The variable *salary* is the CEO salary of the firm i, *mktval* is market value of the firm, *promarg* is profit as a percentage of sales, *ceoten* is years as CEO with the current company, and *comten* is total years with the company. Suppose you want to check if Assumption (1) holds.

(a) (10 points) Derive a test statistic for the null hypothesis $H_0 : \Pr[E(y_i|X_i) = X_i'\boldsymbol{\beta}] = 1$ for some $\boldsymbol{\beta}$, against the alternative hypothesis $H_1 : \Pr[E(y_i|X_i) \neq X_i'\boldsymbol{\beta}] > 0$ for all $\boldsymbol{\beta}$. Note that the null hypothesis $H_0 : \Pr[E(u_i|X_i) = \mathbf{0}] = 1$ implies that $\Pr[Eu_i h(X_i) = \mathbf{0}] = 1$ for any (measurable) function $h(\cdot)$. Construct a test to check this with choosing $h(X_i) = (ceoten_i^2 \ comtem_i^2)'$.

(b) (10 points) Show that the statistic you constructed in (a) can be written in the asymptotically equivalent form, nR_u^2, under Assumption (5), where R_u^2 is the R-squared obtained from regressing $\hat{u}_i$ on X_i, $ceoten_i^2$, and $comtem_i^2$.

5. (20 points) Suppose we run the following two OLS regressions to explain major league baseball players' salaries:

$$\begin{aligned} \log(\widehat{salary}) &= \underset{(.29)}{11.10} + \underset{(.0121)}{.0689}\ years + \underset{(.0026)}{.0126}\ gamesyr + \underset{(.00110)}{.00098}\ bavg + \underset{(.0161)}{.0144}\ hrunsyr + \underset{(.0072)}{.0108}\ rbisyr \\ n &= 353, SSR = 183.186, R^2 = .6278 \end{aligned}$$

and

$$\begin{aligned} \log(\widehat{salary}) &= \underset{(.11)}{11.22} + \underset{(.0125)}{.0713}\ years + \underset{(.0013)}{.0202}\ gamesyr \\ n &= 353, SSR = 198.311, R^2 = .5971 \end{aligned}$$

where *salary* is the 1993 total salary, *years* is years in the league, *gamesyr* is average games played per year, *bavg* is career batting average, *hrunsyr* is home runs per year, and *rbisyr* is runs batted in per year.

(a) (5 points) Suppose we want to test a joint null hypothesis that, once years in the league and games per year have been controlled for, the three performance measures of players – *bavg*, *hrunsyr*, and *rbisyr* – have no effect on salary. The joint hypothesis may be written as $H_0 : \beta_3 = \beta_4 = \beta_5 = 0$. State the alternative hypothesis. Compute a statistic. What is your conclusion? State all the assumptions you need to make. Do not make any assumption you do not need.

(b) (10 points) Suppose you suspect that Assumption (5) may fail. In this case, how would you modify the test statistic in (a) to test for $H_0 : \beta_3 = 0, \beta_4 = 0, \beta_5 = 0$?

(c) (5 points) How would you test whether Assumption (5) holds?