ECO 5350   Prof. Tom Fomby
Introductory Econometrics   Fall 2008

TOPICS TO CONSIDER
FOR MID-TERM

    The Mid-term in this class is scheduled for Wednesday, October 22, 2008 at the regular class time. Don't spend a whole lot of time memorizing formulas. For the exam I will make available to you the formulas on the front and back covers of your Hill, et. al. textbook. Given that we will not have access to a computer during the test you should review up on how to determine critical values for hypothesis tests directly from statistical tables like the Z, t, and F tables. You should commit to memory the ANOVA table I presented in class.

    Our goal before the test is to finish three exercises, **Exercise 3** (test of a single linear hypothesis), **Exercise 4** (test of multiple linear hypotheses), and **Exercise 5** (prediction and prediction intervals in linear regression). When we return from fall break we can talk about when we will hold a review "session."

1. The test will cover the material in Chapters 1 – 6 in your Principles of Econometrics textbook. Also, roughly, this covers the material in the Lecture Notes 1 – 12, 13 (testing a linear restriction), 14 (testing multiple linear restrictions), and 18 (prediction in using linear regression) posted on the course website. Of course you should review the QQs and Exercise Keys posted on the class website.
2. What are the different types of data sets that economists are called on to investigate? **Cross-Section data**, **Time Series data**, and **Pooled Times Series**, and **Panel data.** What distinguishes these different data sets? Can you give me an example of each? Why is it important to be able to distinguish between them?
3. How does one go about testing for the equality of the means of two populations? What is the difference between the Exact t-test and the Approximate t-test? What assumption represents the distinction of the two cases? What is the Folded F-test used for? What is the null hypothesis of the test? What is the alternative hypothesis of the test?
4. One should know the difference between one-sided (one-tailed) versus two-sided (two-tailed) hypothesis tests and how they arise in practice. Correspondingly, one should know the difference between one-tailed and two-tailed p-values and how they are constructed.
5. What is the distinction between the **Classical and Bayesian points of view** of statistics?
6. Appendix C of your textbook provides a very nice review of statistical inference. You need to read through it.
7. What are the **two basic parts** of a SAS computer program? When you run a SAS program, what are the two files that are created? What is a **program file**? A **log file**? An **output (listing) file**? What kind of information do they contain?
8. You should be able to list and understand the Assumptions of the Simple Linear Regression model. See pages 14 and 16 in your textbook.
9. What does the term **homoskedasticity** mean? Draw a population regression function ($E(x \mid y)$) and several conditional density functions at various values of x while at the same time demonstrating the phenomenon of homoskedasticity. Do the same except draw a population regression and several

conditional density functions at various values of x that demonstrate the phenomenon of **heteroskedasticity**. In looking at a data scatter of observations of y and x, how would you determine whether the population regression function is subject to homoskedasticity or heteroskedasticity? Is it OK to use OLS when you have heteroskedasticity in the errors of your regression model?

10. The **Ordinary Least Squares** estimators of the coefficients β_1 and β_2 in the simple linear regression model are derived by using the calculus. What function do you minimize in order to get the OLS estimators? The nice thing about the least squares criterion is that it gives rise to **analytic solutions** for the coefficient estimators $\hat{\beta}_1$ and $\hat{\beta}_2$.

11. Define the term **Population Regression Function (PRF)**. What is a **Sample Regression Function (SRF)** and how does it compare to the population regression (conditional mean) function? What is meant by the "**repeated sampling view**" of statistical hypothesis testing? In words, explain what is meant by "the **sampling distribution** of $\hat{\beta}_2$?" What is meant when we say "$\hat{\beta}_2$ is an **unbiased estimator** of β_2?"

12. What is the meaning of the term "perfect collinearity?"

13. What are the **properties of the Ordinary Least Squares estimator $\hat{\beta}_2$**? What are the properties of the Ordinary Least Squares estimator $\hat{\beta}_2$? What does it mean for these estimators to be **BLUE**? In other words, what does the **Gauss-Markov theorem** state? Can you give me a drawing that represents the Gauss-Markov Theorem? What is the **Extreme Value estimator**? How do you use it to estimate the intercept β_1 and the slope β_2 of the conditional mean function (E(y|x)? Is the Extreme Value estimator unbiased? How does its sampling variance compare to the corresponding Ordinary Least Squares estimators? What does this have to say about the Gauss- Markov Theorem?

14. Hypothesis Testing. What additional assumption on the error term e allows us to conduct small sample inference in the simple linear regression model?

15. How do you display an estimated regression function when presenting it to someone else for his/her inspection? (I.e. reporting coefficient estimates, standard errors and the like). See section 4.2.4 in your textbook.

16. How do you construct a **95% confidence interval for** β_2 using $\hat{\beta}_2$ and se($\hat{\beta}_2$)? How do you use such a confidence interval to test the null hypothesis that $H_0 : \beta_2 = 0$ versus the alternative hypothesis $H_1 : \beta_2 \neq 0$? How do you construct a t-test of the above hypothesis?

14. What is the difference between testing **a two-sided alternative** and testing **a one-sided alternative**? In a "word problem" would you be able to distinguish between when you would use a two-side test versus using a one-sided test? Can you get the p-value of an observed t-statistic for a two-sided alternative? A one-sided alternative?

17. In the **sum of squares decomposition**, TSS = ESS + RSS, what is TSS, ESS, and RSS? What are their mathematical formulas?

18. What is the **coefficient of determination** (R^2)? How do you interpret $R^2 = 0.90$?

19. What is the test for **Overall Significance** of the Regression equation? How does this relate to the **ANOVA table**?

20. A **Test of a Single Linear Hypothesis**. See Lecture notes 13 and Vote.sas on the class website that we discussed in class. Also see Exercise 3 on the Cobb-Douglas production function (Exercise 6.11 in your textbook).

21. **Tests of Multiple Linear Hypotheses in Linear Regression**. See Lecture notes 14 and Exercise 4 on the class website (Exercise 6.9 in your textbook).

22. **Prediction in regression**. There are three types of prediction problems: (1) Prediction of the mean of Y given a certain setting of $X = X_0$, (2) Prediction of the next value of Y given a

certain setting of X = X_0, and (3) Prediction of the next value of Y given an uncertain setting of X around $X = X_0$. See Lecture Notes 18 and Exercise 5.

23. **The Additive/Multiplicative Dummy Variable form of a regression equation.** How do you form the Additive/Multiplicative Dummy Variable form of a regression equation? What is a **dummy variable**? What is the **base (reference) case**? What is the interpretation of the coefficient on the additive dummy variable? What are the interpretations on the multiplicative dummy variables?

24. **The Chow Test for Structural Difference in a Regression Equation.** What is the null hypothesis for the Chow test? What is the alternative for the Chow test? You should know how to use the Additive/Multiplicative Dummy Variable form of a regression equation to conduct a test for structural difference in a cross-section data set.

25. **What are Residual Plots?** How can they be used to examine for potential heteroskedasticity in a regression model?

26. **What is White's test for heteroskedasticity?** What is the null hypothesis for this test? What is the alternative hypothesis of this test? What is the form of the auxiliary equation for the test? How does the overall F-test in this auxiliary equation play into the execution of White's test?