Final Review
ECON 210 Econometrics A

Evan Zuofu Liao*

December 3, 2015

This is a review sheet that outlines the major topics and concepts covered in this course. This review is not meant to be all inclusive, but hopefully it helps you study some of the key points that are essential for the final exam.

1 Statistics

1.1 Probability Basics

- **Random Variables and Probability Distributions**: What is a random variable? What is the difference between discrete and continuous random variables? What are the PMF/PDF and the CDF of a random variable?

- **Expected Values and Variance** What are the properties of expectation? Know how to work with the Law of Iterated Expectations. What are the properties of variance? What are the higher order moments?

- **Joint and Marginal Distributions** How are the the joint, marginal and conditional distributions related to each other?

- **Independence and Correlation** What is the definition for independence? What about mean independence? Does independence imply correlation, and vice versa? If two random variables are independent, what can we say about the expected value and the variance of their product?

- **Common Distributions** Know the properties of normal distribution, t distribution, Chi-square distribution and F distribution.

*Comments and questions to evanliao@uchicago.edu.

1.2 Estimation

- **Properties of Estimators**: What is an estimator? What are the definitions of unbiasedness, efficiency and consistency? Know how to prove these properties using some of the theorems below. Get familiar with the tips and tricks for computing the expected value and the variance of an estimator.

- **Large Sample Theory**: convergence in probability vs. convergence in distribution. What is a sampling distribution of the sample average? What is the Weak Law of Large Numbers? What is the Central Limit Theorem? Know how to use the Continuous Mapping Theorem and the Slutsky Theorem.

1.3 Inference

- **Asymptotic Distribution**: Know how to use the CLT to derive the asymptotic distribution of an estimator (or a test statistic).

- **Hypothesis Testing**: What are the steps in hypothesis testing?

    - Form the null hypothesis
    - Create the test statistic based on your estimator
    - Derive the asymptotic distribution of the test statistic
    - Use the asymptotic distribution to determine the critical value, p-value, confidence interval, etc.

- **Other Things in Testing** : one-sided vs two-sided tests, Type I and Type II error, the power of a test

2 OLS

2.1 Three Interpretations of Linear Regression

Which of the following interpretations gives us $\mathbb{E}\left[u|X\right] = 0$, $\mathbb{E}\left[u\right] = 0$ and $\mathbb{E}\left[Xu\right] = 0$?

- **Best Linear Predictor of $\mathbb{E}\left[Y|X\right]$**

- **Forecasting Interpretation**

- **Causal Interpretation**

2.2   Estimation

- **Minimization problem**: Know how to derive the OLS estimator from the minimization problem. What are the moment conditions?

- **OLS Estimators**: For a regression with a single regressor, what are the population estimators for β_1 and β_0? What are the sample estimators (i.e. $\hat{\beta}_1$ and $\hat{\beta}_2$)?

- **Properties of OLS estimator**: Know how to show the sample OLS estimators are consistent and unbiased, and know how to derive the asymptotic distribution. What is the additional assumption we have to make in proving unbiasedness?

- **Regression without an intercept**: Know how to derive the estimators in a regression without an intercept. Understand the interpretations of the model.

2.3   Inference

- **Hypothesis testing and confidence interval**: Know how to use the asymptotic distribution of the sample OLS estimator to do hypothesis testing for a single restriction and derive its confidence interval.

3   OLS for Multiple Regressors

3.1   Estimation

Before you start to derive the estimators, familiarize yourself with basic matrix algebra.

- **OLS Estimators**: What is the OLS estimator in a population regression? What are the sample analogs? Remember that there are two ways to write down the sample OLS estimator (using the vector notation or using the matrix notation).

- **Perfect Collinearity**: What is perfect collinearity? How does it relate to the invertibility of the $(\mathbb{X}'\mathbb{X})$ matrix? How do we detect high degrees of collinearity in practice?

3.2   Hypothesis Testing

- **Single Restriction**: Know how to define the restriction matrix. What test should we use when we have a small sample? What about a large sample?

- **Multiple Restrictions**: Know how to form the restriction matrix for multiple restrictions (e.g., for a joint significance test). What test should we use when we have a small sample? What are the different implementations of the F-test? What should we do if we don't know the variance-covariance matrix?

3.3   Other Topics in OLS

- **Homoskedasticity**: What is homoskedasticity? What is the asymptotic variance of the OLS estimator under homoskedasticity?

- **Goodness of fit**: What is R^2? How does it relate to the adjusted R^2? What are the TSS, ESS and SSR of a regression?

- **Standard error vs standard deviation**: What are the standard errors of the OLS coefficients? How does it relate to the standard deviation? where do we use standard errors in hypothesis testing?

3.4   OLS in Practice

- **Interpretation**: How does the interpretation of the coefficients change when it is a multiple regression as opposed to a single regression?

- **Log transformation**: What are the interpretations of the coefficients when you take log on the regressor only, on the dependent variable only, and on both sides?

- **Interaction terms**: What is the interpretation for the interaction term? How does the interpretation of other coefficients change when the interaction is added?

- **Higher order terms**: Why do we need higher order terms?

- **Dummy variable regression**: Be sure you know how to derive the OLS coefficients in a regression run on multiple dummy variables. Know the interpretation of the coefficients when we have an interaction of two dummy variables.

- **The linear probability model**: What is a linear probability model? How to interpret the coefficients when we have a linear probability model?

- **Threats to internal validity**: Understand why and how the following problems may affect our estimation of the coefficients: omitted variable bias, measurement error, sample selection bias, simultaneous causality bias

- **Data generating process vs data observed**: What is the difference between these two? Which model, the true model or the observed model, should we plug in to derive the bias arising from measurement error and omitted variable? What is attenuation bias? What is a classic measurement error?

4 IV

4.1 Motivation

- **Endogeneity**: Why do we need IV? What are the assumptions we need when we are running an IV regression? What is an endogenous variable as opposed to exogenous variable?

- **Interpretation**: What is the story you tell when you run an IV? What is the interpretation of the IV coefficient?

4.2 Estimation

For the following cases, you should know how to derive the estimators and do hypothesis testing

- **IV with a single regressor and a single instrument**: IV estimator

- **IV with multiple regressors and multiple instruments (exactly identified, # of instruments = # of endogenous regressors)**: IV estimator

- **IV with multiple regressors and multiple instruments (over identified, # of instruments > # of endogenous regressors)**: 2SLS estimator

4.3 Instrument Validity

- **Two conditions**: what are the two conditions the instruments have to satisfy?

- **Weak instrument**: what is a weak instrument? How do we carry out a weak instrument test?