

EE 325 Introductory Econometrics

By Wanwiphang Manachotphong¹

Semester 2, 2016

¹Email: wanwiphang@econ.tu.ac.th; Office: 525 Faculty of Economics, Tha Prachan; Office hours: Tuesdays, 11.00 – 12 noon (or by appointment)

Contents

1	Introduction and Motivation	3
1	What is econometrics?	3
1.1	Steps in Empirical Economics Analysis	5
2	Types of economic data	6
2	Review of Some Statistical Concepts	7
1	Random variables and distributions	7
1.1	Types of Random Variables	8
1.2	Distribution of a Random Variable (X)	9
1.3	Examples of Some Distributions	10
1.4	Population Distribution vs. Sample Distribution	11
2	Joint Distributions, Conditional Distributions and Independence	12
2.1	Joint Distributions	12
2.2	Conditional Distribution and Marginal Distribution	13
2.3	Independence	16
3	Expectation, variance, covariance and correlation	17
3.1	Expected value of a Random Variable – $E(X)$ or μ_X	17
3.2	Properties of Expected Values	18
3.3	Conditional and Marginal Expectations	19
3.4	Variance and Standard Deviation of a Random Variable	20
3.5	Properties of Variances	21
3.6	Properties of Standard Deviations	21
3.7	Covariance and Correlation of Two Variables	23
4	Estimators and desirable properties of estimators	25
4.1	Desirable properties of estimators	26
4.2	The concept of Mean Squared Error (MSE)	29
3	The Simple Regression Model	31
1	Principle, assumptions and derivation of ordinary least squares (OLS) estimators	32
1.1	Terminology for the Linear Regression	32
1.2	Derivation of Ordinary Least Squares (OLS) Estimators	33
2	Properties of OLS estimators	36
2.1	Algebraic Properties	36
2.2	Properties proving BLUE	37
2.3	Assumptions on the simple linear regression (SLR) model	37
2.4	Homoskedasticity VS. Heteroskedasticity	38
2.5	Expectation of Estimators	39
2.6	Variance of OLS Estimators	40
2.7	Some Concepts to be exphasized (Population vs. Sample; PRF vs. SRF; error vs. residual)	41

2.8	Goodness of Fit (R^2)	43
2.9	Incorporating Nonlinearities in Sample Regression	44
3	Regression Through the Origin	45
4	Multiple Regression Analysis (Estimation)	47
1	Motivation	47
1.1	Assumption SLR 4 ($E(u X) = 0$) in the Multiple Regression Context	48
2	The Model with k Independent Variables	49
2.1	Accounting Nonlinearity	50
3	Estimation of parameters and properties of estimators	51
3.1	Deriving OLS Estimators	51
3.2	Algebraic Properties of OLS estimators	51
3.3	How could the multiple regression analysis enable ceteris paribus analysis?	52
4	Expected Value of the OLS Estimators	53
4.1	Issue #1: Including Irrelevant Variable (Overspecifying the Model)	53
4.2	Issue #2: Excluding Relevant Variable (Underspecifying the Model → omitted variable bias. This is a serious problem!)	54
5	Variance of the OLS Estimators	55
6	Estimator of the OLS Variance	56
5	Stata Lab 1 – Introduction	57
1	What is STATA?	57
1.1	STATA supports	57
1.2	Data files and Do-files	57
2	Tutorial 1: Exploring the Data and Running a Simple Regression	58
2.1	Examples from Wooldridge(2009)	60
6	Multiple Regression Analysis (Inference)	61
1	Sampling Distribution of the OLS estimators ($\hat{\beta}_{OLS}$)	62
2	Testing Hypotheses about an individual regression coefficient "the t-test"	63
2.1	Testing Against Two-Sided Alternatives	63
2.2	Testing Against One-Sided Alternatives	65
3	Testing other hypotheses about β_j	66
4	Testing Hypotheses about a Single Linear Combination of the Parameter	67
5	Computing p-Values for t-Tests	68
6	Confidence Intervals (CI)	69
7	Testing Multiple Linear Restrictions: The F-test	71
8	How the Hypothesis Testing is done in Practice	74
7	Multiple Regression Analysis : Further Issues	75
1	Data scaling on OLS statistics	75
2	More on functional forms	76
3	Models with Interaction Terms	78
4	More on the Goodness-of-Fit and Selection of Regressors	79
8	Multiple Regression Analysis with Qualitative Information:	81
1	Outline	81
2	Describing Qualitative Information	81

3	Models with a single dummy independent variable	82
4	It is not possible to include all of the dummy alternatives in the same model	83
5	Using dummy variables for multiple categories	84
6	Interactions involving dummy variables	87
7	Testing for Differences in Regression Functions across Groups	89
	7.1 We can use the "Chow statistics" to test this type of hypothesis as well	90
8	A Binary Dependent Variable (y variable): The Linear Probability Model	92
9	Heteroscedasticity Problem	95
1	Nature and Consequences of heteroscedasticity for OLS	95
	1.1 Nature of Heteroskedasticity	95
	1.2 Consequences of Heteroskedasticity	95
	1.3 How can the estimated value of $Var(\hat{\beta}_{OLS})$ be wrong?	96
	1.4 Two types of remedies	97
2	Testing for heteroskedasticity	98
	2.1 Breusch-Pagan test (BP test)	99
	2.2 The White Test	102
3	Remedial measures	103
	3.1 Weighted Least Squares (WLS)	103
	3.2 Feasible GLS	104
	3.3 What if the assumed h_i function is wrong?	105
10	Lab 2 – Dummy, Heteroskedasticity, Specification Issues	107
1	Does "beauty" help increase wage?	107
2	Fixing Heteroskedasticity	109
3	Labor Force Participation of Female	110
11	Serial Correlation and Heteroskedasticity in Time Series Regressions	113
1	The Nature of Time Series Data	113
2	Examples of Time Series Regression Models	114
	2.1 Static Models	114
	2.2 Finite Distributed Lag Models	114
3	Properties of OLS under classical assumptions	115
4	Properties of OLS with Serially Correlated Errors	117
5	Unbiasedness and Consistency	117
6	Efficiency and Inference	117
7	Testing for Serial Correlation	118
	7.1 A "t-test" for AR(1) serial correlation with strictly exogenous regressors	118
	7.2 The Durbin-Watson Test (DW test)	119
	7.3 Testing for AR(1) serial correlation "without" strictly exogenous regressors	120
	7.4 Testing for AR(q) serial correlation "without" strictly exogenous regressors	121
8	Correcting for serial correlation	121
	8.1 Passive way	121
	8.2 Active way –	121

12 Multicollinearity	123
1 The Nature of Multicollinearity	123
2 Consequences of Multicollinearity	123
2.1 The OLS estimator will still be BLUE.	123
2.2 The variances and covariances will be very large. This makes precise estimation difficult.	123
3 Detection of multicollinearity	124
4 Remedial Measures	125