

## Learning Goals

### Grand Objectives

- **Be able to understand, interpret, and implement multiple regression and related statistical techniques**
- **Know the limitations and pitfalls of regression methods**
- **Be able write a focused explanation of the findings of a statistical investigation, clearly and concisely and “very well written.”**

### Specific Objectives

#### 1. Statistics Review for regression

- 1.1. Light review of statistics
  - 1.1.1. Probability distributions
    - 1.1.1.1.Expected value
    - 1.1.1.2.Variance and Standard Deviation
    - 1.1.1.3.Covariance and correlation
  - 1.1.2. Population versus sample statistics
    - 1.1.2.1.Algebra of summations and sums of squares
  - 1.1.3.
- 1.2. Interpreting coefficients as *ceteris paribus* marginal responses.
  - 1.2.1. Familiarity with both simple and multiple regression
- 1.3. Review of statistical properties of an estimator
  - 1.3.1. Idea of a Data Generating Process (DGP)

1.3.2. Distribution of estimator over repeated draws from a DGP

1.3.2.1. Mean (Is it Unbiased?) and Variance

1.4. Hypothesis tests on a simple mean

1.4.1. One-tailed and two-tailed  $t$ -tests

1.4.2. Confidence interval

## 2. Multiple Regression in the Classical Case

2.1. Least squares as data-fitting technique

2.1.1. Simple regression without constant

2.1.2. Simple regression with constant and intuition for  $k$ -RHS variable model

2.2. Least squares parameters as statistical estimators

2.2.1. Derive mean and variance of  $\hat{\beta}_{ols}$ .

2.2.1.1. Being able to use simple statistical algebra

2.2.1.2. Understanding relevance of classical OLS assumptions

2.2.2. Understand determinants of variance of  $\hat{\beta}_i$

2.2.2.1. 1 RHS variable regression

2.2.2.2. 2 RHS variable regression

2.2.2.3. Multicollinearity

2.2.3. Prove Gauss-Markov theorem

2.3. Hypothesis testing and confidence intervals

2.3.1. Wald and likelihood-ratio tests

2.3.2.  $t$ -tests

2.3.3.  $F$ -tests

2.3.3.1. likelihood-ratio version

2.3.3.2. Chow test

2.3.4. Confidence intervals and regions

2.3.4.1. 1 RHS variable confidence interval

2.3.4.2. Joint confidence ellipse for 2-RHS variables

2.3.5. Type 1 versus Type 2 error

2.3.6. Size and Power

2.4. Data errors and outliers

2.5. Omitted variables bias

- 2.6. functional forms
  - 2.6.1. Cross section models
    - 2.6.1.1. Dummy variables
    - 2.6.1.2. logs and other nonlinear transformations
  - 2.6.2. Time series models
    - 2.6.2.1. Lagged exogenous variables
    - 2.6.2.2. Lagged endogenous variables
- 2.7. Forecasting

### **3. Generalized least squares**

- 3.1. Known form of heteroskedasticity
  - 3.1.1. Impact on OLS
  - 3.1.2. GLS estimation
- 3.2. Heteroskedasticity
  - 3.2.1. Models of heteroskedasticity
- 3.3. Serial correlation
  - 3.3.1. testing
    - 3.3.1.1. Durbin-Watson
    - 3.3.1.2. Breusch-Godfrey
  - 3.3.2. estimation
    - 3.3.2.1. Cochrane-Orcutt

### **4. Alternatives to least squares – of which we’ll discuss some**

- 4.1. Instrumental variable estimation
- 4.2. Errors-in-variables
- 4.3. Simultaneous equation models
- 4.4. Panel data
- 4.5. Maximum likelihood estimation
- 4.6. Nonlinear least squares
- 4.7. Probit/Logit

### **5. Computer skills**

#### **5.1. EViews**

### **6. Writing Skills**

- 6.1. Focused writing which identifies a narrow question and answers it
  - 6.1.1. Separation of writing and editing process
  - 6.1.2. Writing for a specific audience

### 6.1.3. Rewriting

#### 6.1.3.1. Rewriting

##### 6.1.3.1.1. Rewriting