

Economic Data Analysis Using R

- Introduction to R
 - Getting Started
 - Using RStudio IDE
 - R Basics
- Economic Data
 - Cross Sections
 - Time Series
 - Panel Data

Introduction to R

- Getting Started
 - [R 3.2.x](#)
 - [RStudio 0.99.xxx](#)
- Using Rstudio IDE
 - Console (Input/Output)
 - Editor (Script, Data, Project)
 - History (Environment, Search)
 - Help (Viewer, Files, Plots, Packages)

Getting Started

- R Basics
 - [Roger D. Peng: Computing for Data Science](#)
 - Data Types and Basic Operations:
[Note 1](#), [Note 2](#), [Note 3](#), [Note 4](#)
 - Reading and Writing Data:
[Note 5](#), [Note 6](#)
 - Functions:
[Note 7](#)

Introduction to R

- R Resources

- [R-Projects \(CRAN\)](#)
- [R-bloggers](#)
- [Quick-R](#)
- [DataCamp](#)

- R References

- [An Introduction to R](#), by W. N. Venables, D. M. Smith and the R Core Team, 2015
- Roger D. Peng, [R Programming for Data Sciences](#), Learnpub.com, 2015

Economic Data

- Data Structure
 - Cross Sections
 - Time Series
 - Panel Data
- Random Experiments
 - $x_i \sim iid \text{ (normal)}$
 - $x_t \sim arima$
 - $x_{it} \sim iid + arima$
- Data Size
 - High Frequency Financial Data
 - Massive Datasets (Big Data)
 - Public Administrative Data
 - Private Company Data

Economic Data

- Data Generating Process (DGP)

- Economic Theory

- Probability Distribution

uniform $f(x; \min, \max) = \frac{1}{\max - \min}, \min \leq x \leq \max$

normal $f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right), -\infty < x < \infty$

binomial $f(x; n, p) = \binom{n}{x} p^x (1 - p)^{n-x}, x = 0, 1, \dots, n$

Poisson $f(x; \lambda) = \lambda^x \frac{\exp(-\lambda)}{x!}, x = 0, 1, 2, \dots$

...

Economic Data

- High Frequency Financial Data
 - Stock Market Daily Time Series
 - S&P 500, VIX (Fear Index), China SSEC
 - Exchange Rates
 - Ticker Data
 - Apple Co. (AAPL)
 - Tesla Motors, Inc. (TSLA)

Economic Data

- Useful On Line Data Resources
 - [Quandl](#)
 - [quantmod](#)
 - [Kaggle Datasets](#)
 - [Google Public Data](#)
 - [Amazon AWS Public Data](#)
 - [Intraday Stock Quotes](#)

Economic Data Analysis

- Data Preparation
 - Getting Data
 - Tidying Data
- Data Visualization
 - Using Graphs
- Data Analysis
 - Regression
 - Classification

Data Preparation

- Getting Data
 - Using R: `read.table()`, `read.csv()`, `read.dta()`, `read_excel()`
- Tidying Data
 - Structuring datasets to facilitate visualization and analysis:
 - Each variable forms a column; Each observation forms a row; Each type of observational unit forms a table.
 - Using R: `reshape()`, `subset()`, `order()`, `transform()`, `aggregate()`

Data Visualization

- Simple Graphs
 - Scatter, Line, Bar, and Box Plots
 - Histogram and Distribution
- Using ggplot2 Package
 - Grammar of Graphics ([Part1](#), [Part2](#))
- References
 - [Data Visualization with ggplot2](#)
 - [Tutorial Example](#)

R Package ggplot2: qplot

- `qplot(data, x, y, ..., geom(...)) + ...`
 - Data Frame
 - Aesthetics: `x`, `y`, `alpha`, `color`, `size`, `shape`, `fill`, ...
 - Geometry: `point`, `line`, `bar`, `jitter`, `boxplot`, `histogram`, `density`, `smooth`, ...
- Adding facets
 - `facet_grid`

R Package ggplot2: ggplot

- `ggplot(data, aes(x,y,...), geom_...) + ...`
 - Data
 - Aesthetics
 - Geometry
 - Stats, Facets
 - Scales, Colors, Coordinates
 - Labels, Legends
 - Themes

Data Analysis I

- Descriptive Statistics
- Correlation and Covariance
- Analysis of Variances (AOV, ANOVA)
- Hypothesis Testing
 - DGP $\sim$ Non IID
 - One-Variable Testing (t-test)
 - Two-Variable Testing (paired t-test)

Data Analysis I

- Cross Sections Data
 - Hypothesis Testing
 - Normality
 - Independence
 - Homoscedasticity
 - Linear Regression
 - Least Squares Regression
 - Maximum Likelihood

Data Analysis I

- Time Series Data
 - Hypothesis Testing
 - Durbin-Watson
 - Box-Pierce / Ljung-Box
 - ACF/PACF
 - Transformation: Lag, Difference
 - Model Estimation
 - Forecasting

Data Analysis I

- Panel Data
 - Hypothesis Testing
 - Unobserved Hetrogeneity
 - Spatial Correlation
 - Model Estimation
 - Fixed Effects
 - Random Effects
 - Multilevel Analysis