

## Syllabus

# Research Methods in Political Science I

School of Law, Kobe University, Fall 2015

Wed. 5:00–6:30pm  
Seminar Room, Computer Lab, Building No.3  
(第三学舎 情報処理演習室)  
Office Hours: Mon & Wed 12:00–1:00pm  
(and by appointment)

Instructor : Yuki Yanai (矢内 勇生)  
Email: [yanai@lion.kobe-u.ac.jp](mailto:yanai@lion.kobe-u.ac.jp)  
Website: <http://yukiyanai.com>  
Office: Room 404, Building No.4  
(第四学舎 404 研究室)

### Overview and Class Goals

This is a course on quantitative political methodology, which utilizes statistical methods to investigate problems in political science. You will learn to conduct reproducible research in political science with quantitative data. More specifically, you will learn:

- how to find data that enable you to answer your research questions,
- how to transform the collected data so that you can efficiently analyze them,
- how to analyze the data,
- how to report the result of your analyses, and
- how to record your entire data-analysis projects.

### Prerequisites

There is no official prerequisites for this course, but basic understanding of statistics (up to linear regression) is essential. You might want to take at least one undergraduate class in statistics or quantitative methods before taking this course. If you are not confident in your knowledge about statistics, please consult the following books.

- Rowntree, Derek. 2000. *Statistics without Tears*. London: Penguin. (加納悟 訳. 2001. 『新・涙なしの統計学』 新世社.)
- Moore, David S., George P. McCabe, and Bruce A Craig. 2014. *Introduction to the Practice of Statistics*, 8th ed. New York: W.H. Freeman.
- Blitzstein, Joseph, and Jessica Hwang. 2014. *Introduction to Probability*. Boca Raton: CRC Press.
- 小島寛之. 2006. 『完全独習 統計学入門』 ダイアモンド社.
- 山田剛史, 杉澤武俊, 村井潤一郎. 2008. 『Rによるやさしい統計学』 オーム社.

### Class Format

Class meets in a computer lab. A typical lesson consists of a 60-minute lecture and a 30-minute lab session. Students can use an iMac (with a student ID and the password) or your own laptop computer. *You are not allowed to eat or drink in the lab.*

## Requirements

Grades will be based on

- class participation (20% of final grade),
- homework assignments (50%), and
- final project (30%).

The weekly homework assignments will consist of analytical problems, computer simulations, and data analysis. They will usually be assigned right after a class meeting and due the following Wednesday morning (9am). *No late submission will be accepted.* Detailed explanation about the final project will be given in class.

## Course Website

The course website is located at the following URL:

<http://www2.kobe-u.ac.jp/~yyanai/classes/rm1/contents/>

You are expected to check the website on regular basis (please refresh your web browser to view the latest content). The weekly assignment will be posted on the website.

## Computation

You will learn R, an open-source computing language, to collect, manage, and analyze data. I will teach how to use R in RStudio, an integrated development environment (IDE) for R. For more information about R and RStudio, visit the course website.

If you prefer to use another statistical software packages (e.g., Stata or SPSS), you can use it for the course at your own risk.

## Required Book

Please purchase the following book. Either edition (English or Japanese) is fine.

- Lander, Jared P. 2014. *R for Everyone: Advanced Analytics and Graphics*. Upper Saddle River: Addison-Wesley. (高柳慎一ほか訳. 2015. 『みんなの R : データ分析と統計解析の新しい教科書』マイナビ.)

## Optional Books

The following books are optional. Many of you should find them useful, though.

1. Grolemund, Garret. 2014. *Hands-On Programming with R*. Sebastopol: O'Reilly. (大橋真也監訳. 2015. 『RStudio ではじめる R プログラミング入門』オライリー・ジャパン.)
2. King, Gary. 1998. *Unifying Political Methodology*. Ann Arbor: University of Michigan Press.
3. Gelman, Andrew, and Jennifer Hill. 2007. *Data Analysis Using Regression and Multi-level/Hierarchical Models*. New York: Cambridge University Press.
4. 浅野正彦, 矢内勇生. 2013. 『Stata による計量政治学』オーム社.

## Schedule

The readings with **M** (Mandatory) should be completed prior to the lecture for which they are listed. Graduate students should at least skim readings with **R** (Recommended) as well either before or after the lecture. Readings with **O** (Optional) should enrich your understanding of the topics. *This schedule is subject to change.*

### 1. Introduction (Oct. 7)

- Course Overview
- Introduction to R and RStudio

**M** Lander, chs.1–4 (read after class)

### 2. Statistical Computing with R (Oct. 14)

- Review basics of statistics
- Compute statistics with R

**M** Lander, ch. 4, 6, 8, 14, and 15.

**O** Review statistics by reading a book(s) listed above in Prerequisites section

### 3. Reproducible Research: Introduction to Literate Programming (Oct. 21)

- Reproducible research
- Literate programming
- Markdown and R Markdown

**M** Lander, ch.23

**M** Donoho, David L. “An Invitation to Reproducible Computational Research.” *Biostatistics* 11(3): 385–388.

**R** Beck, Nathaniel. 2014. “Research Replication in Social Science.” *OUPblog*. 24 August 2014.

**R** King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press (真淵勝 監訳. 2004. 『社会科学のリサーチ・デザイン：定性的研究における科学的推論』勁草書房) ch.1

**R** King, Gary. 1995. “Replication, Replication.” *PS: Political Science & Politics* 28: 444–452.

**O** Gandrud, Christopher. 2015. *Reproducible Research with R and RStudio*, 2nd ed. Boca Raton: CRC Press.

**O** Xie, Yihui. 2013. *Dynamic Documents with R and knitr*. Boca Raton: CRC Press.

**O** Freese, Jeremy. 2007. “Replication Standards for Quantitative Social Science: Why Not Sociology?” *Sociological Methods & Research* 36(2): 153–172.

### 4. Visualizing Data and Results: Tables and Figures (Oct. 28)

- Visualizing data
- Creating tables and figures in R

**M** Lander, ch.7

- M Kastellec, Jonathan P., and Eduardo L. Leoni. 2007. “Using Graphs Instead of Tables in Political Science.” *Perspectives on Politics* 4: 755–771.
- R Gelman, Andrew. 2011. “Why Tables Are Really Much Better Than Graphs.” *Journal of Computational and Graphical Statistics* 20(1): 3–7. (A working paper version is available here.)
- O Wickham, Hadley. 2010. *ggplot2: Elegant Graphics for Data Analysis*.
- O Murrel, Paul. 2006. *R Graphics*. Boca Raton: Chapman & Hall/CRC. (久保拓弥 訳. 2009. 『R グラフィクス : R で思いどおりのグラフを作図するために』 共立出版)
- O Tufte, Edward. 2001. *The Visual Display of Quantitative Information*. Cheshire: Graphics Press.
- O Yau, Nathan. 2011. *Visualize This: The FlowingData Guide to Design, Visualization, and Statistics*. Indianapolis: Wiley.
- O Tukey, John W. 1977. *Exploratory Data Analysis*. Reading: Addison-Wesley.

## 5. Collecting Data: Introduction to Web Scraping (Nov. 4)

- Data collection
  - Web scraping
- R Munzert, Simon, Christian Rubba, Peter Meißner, and Dominic Nyhuis. 2015. *Automated Data Collection with R: A Practical Guide to Web Scraping and Text Mining*. Chichester: Wiley.
- O Mitchell, Ryan. 2015. *Web Scraping with Python: Collecting Data from the Modern Web*. Sebastopol: O’Reilly.
- O Downey, Allen B. 2014. *Think Python*. <http://www.greenteapress.com/thinkpython/html/index.html>
- O *Dive Into Python 3*. <http://www.diveintopython3.net/> (日本語版. <http://diveintopython3-ja.rdy.jp/> )
- O Hemenway, Kevin, and Tara Calishain. 2003. *Spidering Hacks: 100 Industrial-Strength Tips & Tools*. Sebastopol, CA: O’Reilly. (村上雅章 訳. 2004. 『Spidering Hacks : ウェブ情報ラクラク取得テクニック 101 選』 オライリー・ジャパン)

## 6. Linear Regression (Nov. 11)

- Review Linear regression
  - Run linear regression in R
- M Lander, ch.16
- M Rohrschneider, Robert, and Stephen Whitefield. 2007. “Representation in New Democracies: Party Stances on European Integration in Post-Communist Eastern Europe.” *Journal of Politics* 69(4): 1133–1146.
- O Fox, John. 1997. *Applied Regression Analysis, Linear Models, and Related Methods*. Thousand Oaks: SAGE Publication.
- O Fox, John, and Harvey Sanford Weisberg. 2011. *An R Companion to Applied Regression*. Thousand Oaks: SAGE Publication.
- O Gelman & Hill, chs.3–4

O 浅野・矢内 第10, 11章

### 7. Linear Regression (cont.) (Nov. 18)

M Lander, ch.18

M Krueger, James S., and Michael S. Lewis-Beck. 2008. "Is OLS Dead?" *Political Methodologist* 15(2): 2–4.

O 浅野・矢内 第12, 13章

### 8. Linear Regression (cont.) (Nov. 25)

M Brambor, Thomas, William Roberts Clark, and Matt Golder. 2006. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 14(1): 63–82.

R Kam, Cindy D., and Robert J. Franzese, Jr. 2007. *Modeling and Interpreting Interactive Hypotheses in Regression Analysis*. Ann Arbor: University of Michigan Press.

### 9. Maximum Likelihood Estimation (Dec. 2)

- Maximum likelihood method

M Lander, ch.17 (only 17.1)

R King, chs.4–5 (**strongly recommended**)

O Kleinbaum, David G., and Mitchel Klein. 2010. *Logistic Regression: A Self-Learning Text. 3rd ed.* New York: Springer.

O Gelman & Hill, ch.5

O Maddala, G. S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*. New York: Cambridge University Press.

O Pawitan, Yudi. 2001. *In All Likelihood: Statistical Modeling and Inference Using Likelihood*. Oxford: Oxford University Press.

O 浅野・矢内 第13章

### 10. Logistic Regression (Dec. 9)

- Analysis of binary outcomes
- Logistic regression

M Gilens, Martin. 2005. "Inequality and Democratic Responsiveness." *Public Opinion Quarterly* 69(5): 778–796.

M Review Lander, 17.1

O Same as Week 9

### 11. Probit Models (Dec. 16)

- Analysis of binary outcomes
- Probit models

M Bartels, Larry M. 2000. “Partisanship and Voting Behavior, 1952–1996.” *American Journal of Political Science* 44(1): 35–50.

R King, ch.5

O Gelman & Hill, ch.6

**No class on Jan. 6.** The instructor is out of town.

## **12. Analysis of Categorical Data (Jan. 13)**

- Ordered logit/probit
- Multinomial logit/probit

M Knack, Stephen. 2004. “Does Foreign Aid Promote Democracy?” *International Studies Quarterly* 48(1): 251–266.

M Dow, Jay K., and James W. Endersby. 2004. “Multinomial Probit and Multinomial Logit: A Comparison of Choice Models for Voting Research.” *Electoral Studies* 23(1): 107–122.

R King, ch.5

O Gelman & Hill, ch.6

O Borooah, Vani K. *Logit and Probit: Ordered and Multinomial Models*. Thousand Oaks: SAGE Publications.

## **13. Analysis of Incomplete Data (Jan. 20)**

- Handling missing data
- Multiple imputation with R

M Ross, Michael. 2006. “Is Democracy Good for the Poor?” *American Journal of Political Science* 50(4): 860–874.

R Su, Yu-Sung, Masanao Yajima, Andrew Gelman, and Jennifer Hill. 2011. “Multiple Imputation with Diagnostics (mi) in R: Opening Windows into the Black Box.” *Journal of Statistical Software* 45

O Little, J. Rodrick, and Donald Rubin. 1987. *Statistical Analysis with Missing Data*. New York: Wiley.

O Rubin, Donald. 1996. “Multiple Imputation after 18+ Years.” *Journal of American Statistical Association* 91: 473–489.

O Schafer, Joseph L. 1999. “Multiple Imputation: A Primer.” *Statistical Methods in Medical Research* 8: 3–15.

O Honaker, James, Gary King, and Matthew Blackwell. 2012. “AMELIA II: A Program for Missing Data.” <http://www.icesi.edu.co/CRAN/web/packages/Amelia/vignettes/amelia.pdf>

## **14. Students’ Presentation (Jan. 27)**

- Interim report of the final project

M King, Gary, Michael Tomz, and Jason Wittenberg. 2000. “Making the Most of Statistical Analyses: Improving Interpretation and Presentation.” *American Journal of Political Science* 44(2): 347–361.

R Leeper, Thomas. 2013. “Making High-Resolution Graphics for Academic Publishing.” *Political Methodologist* 21(1): 2–5.