

Rによる計量分析：データ解析と可視化

経済学特殊講義

富山大学 経済学部 2017年度後期

開講時間:	月曜2限	担当教員:	伊藤 岳
教室:	TBA	Email:	gito@eco.u-toyama.ac.jp
オフィス・アワー:	月曜3限 or 個別相談	研究室:	経済学部研究棟 418号室
Website:	http://cfes-project.eco.u-toyama.ac.jp/education/education_2017/r_2017/		

授業のねらいとカリキュラム上の位置付け (一般学習目標) / Course Objectives

本講義では、基礎的な計量経済・統計学の理論を学んだ(あるいは、これから本格的に学びたい)学生を主な対象として、応用的な計量分析の方法を学ぶ。具体的には、フリーの強力なソフトウェア/統計処理言語であるRを用いたデータの可視化・解析の方法を学び、さらに統計的シミュレーションや現実のデータの解析を通じて統計的推論上の問題(たとえば、欠落変数バイアス)を理解する。あわせて、ときとして煩雑な作業が必要になるデータ取得や統計的シミュレーション、データ解析・可視化の作業をRによって自動化・再現可能にする方法も学び、自らの研究で計量分析を用いる素地を身につける。

達成目標 / Course Goals

この講義では、初歩的な計量経済・統計学の知識を前提に、履修者が次のような応用的な経験・知識・スキルを身につけることを目指す。

- (1) データの可視化・解析のスキルを身に付ける —— 基礎的なデータの可視化および計量経済・統計学の推定を、R言語を用いて行なう方法・スキルを身につける
- (2) 基礎的な推定法、仮定、概念を理解する —— 統計的シミュレーションと経験的な(現実の)データの解析を通して、基礎的な計量経済・統計学の推定法を理解する
- (3) 再現可能な研究を行なう力を身につける —— R言語による実装・プログラミングを通して、再現可能な研究(データの取得・整理・可視化・解析の一連の作業を、他者が追試・確認・再現可能な研究)を行なう方法を身につける

授業計画 (授業の形式、スケジュール等) / Class schedule

Week 1 イントロダクション 10/2

Week 2 Rのセットアップ 10/16

- 第2回以降、個人のラップトップ(ノートPC)持ち込み必須。下記の「授業時間外学修(事前・事後学修)」および「履修上の注意」も確認のこと
- RおよびR Studioのインストール

- パッケージのインストール

Week 3 Rの基礎とデータ操作・管理 (1): データの入出力とプロジェクト 10/23

Week 4 休講 10/30

Week 5 Rの基礎とデータ操作・管理 (2): データの操作と整理 (data tidying) 11/6

Week 6 データの可視化 (1): 基礎的な可視化 11/13

Week 7 データの可視化 (2): 地図, ネットワークを含む発展的な可視化 11/20

- 中間課題 (1) 出題

Week 8 データ取得の自動化: ウェブスクレイピングと API 11/27

Week 9 回帰分析の準備: 基礎概念の確認とシミュレーション 12/4

Week 10 休講 12/11

Week 11 回帰分析 (1): 最小二乗法 (OLS) と内生性の問題 12/18

Week 12 休講 1/9

Week 13 回帰分析 (2): 統計的シミュレーションによる回帰分析の理解 1/15

- 中間課題 (2) 出題

Week 14 回帰分析 (3): 回帰分析の解釈, 限界効果と交互作用項 1/22

Week 15 回帰分析 (4): 一般化線形モデル (GLM) 1/29

Week 16 補講 TBA

- 最終課題出題

授業時間外学修 (事前・事後学修) / Independent Study Outside of Class

事前学習

- 第2回講義までに, 次の事前学習を行なうこと.
 - 次の単語の意味を把握しておくこと (Google や Yahoo! で検索すればすぐに分かる): 「全角文字」「半角文字」「行」「列」「ファイルパス」「ディレクトリ」「拡張子」「文字コード/エンコーディング」
 - 次のような操作/ショートカットキーの表記が, どのようなキーボード操作を意味するかを把握し, 行えるようにしておくこと: “Command + S,” “Command + Shift + K” (macOS), “Control + C,” “Control + Shift + N” (Windows)
 - 各自の OS 環境で任意のファイルのファイルパスを取得する方法を調べ, 実践しておくこと (任意のファイルのファイルパスを取得できるように練習しておくこと)
- 毎回の講義時に指示する, 次回講義に関連する教科書・副読本の章を事前に読んでおくこと

事後学習

- 講義で用いた R コードを確認・(再度) 実行し, 復習すること

これらの事前・事後学修を行なわない場合, 講義の理解や課題提出に支障をきたす. 第2回以降の講義は, 履修者が事前学習を行なっていることを前提に進める. 正当な理由 (たとえばインフルエンザ) がある場合を除き, 欠席者の個別フォローアップは行なわない.

キーワード / Keywords

R, API, ウェブスクレイピング, 可視化, 計量経済学, シミュレーション, データサイエンス, 統計学

履修上の注意 / Notices

- 個人のラップトップ・コンピュータ (ノート PC) を持ち込むこと。毎回の講義では、個人のラップトップで R あるいは R Studio を用いて作業を行なう実習時間を設ける
- R をインストール可能な Linux (Ubuntu, Debian 等), Mac OS, Windows のいずれかであれば OS は問わないが、管理者権限のあるアカウントとパスワードを必ず把握しておくこと。当然ながら、「パスワードが分からないのですが、どうしたらいいですか」等の質問には対応できない
- 事前に学内の無線 LAN に接続できるように設定しておくこと。海外購入のマシンを用いるなど、学内 LAN に接続できない特殊な事情がある場合は、初回講義時までに ITC (情報基盤センター) に相談すること
- R の特性上、この講義ではほぼすべての作業をコード記述・キーボード操作で行なう。クリック操作は基本的に行なわない
- R 言語その他のプログラミング言語についての知識は問わない。経験があれば講義内容や講義で扱う R コードを理解しやすいが、必須ではない
- 初歩的な計量経済・統計学の知識や微積分の基礎を前提とする。計量経済・統計学の知識のない学生の履修も妨げないが、他の科目を履修し推定方法や理論を並行して学ぶことを強く勧める。本講義は厳密な理論的基礎よりも、応用的な力を養うことを重視する
- 使用教室の規定と齟齬が生じない限り、授業中の飲食やスマートフォンの使用は原則的に認める。たとえば、「Google 検索すればわかりそうだけれど、自分の知識では正確に分からない」単語が講義で出た場合には、遠慮なくスマートフォンで検索して構わない。同様に、集中力を維持するために飲食が必要であれば特に問題視しない。ただし、いずれの場合もその行為が適切か否か・正当な理由があるかを考えること (また、他の履修者から苦情があった場合や、頻度過剰と担当教員が判断した場合、自ら説明すること) を求める

教科書・参考書等 / Textbooks

講義資料

毎回の講義で使用するスライドと R コード、サンプル・データ等の資料は、講義用ウェブサイトに前日までにアップロードする。各自ダウンロード／印刷の上、持参すること。紙媒体では配布しない。URL は、http://cfes-project.eco.u-toyama.ac.jp/education/education_2017/r_2017/

教科書 (購入あるいは複写必須)

- Gelman, A & J Hill. 2007. *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge University Press.
- 星野 匡郎・田中 久稔. 2016. 『R による実証分析：回帰分析から因果分析へ』 オーム社.
- Imai, K. 2017. *Quantitative social science: An introduction*. Princeton, NJ: Princeton University Press (近刊の邦訳でも可)
 - － 毎回の講義では、上記の教科書に加えて以下の副読本からも適宜課題文献 (次回講義までに読んでおく章) を割り当てる

副読本 (Rの解説を含むもの・購入推奨)

- 浅野 正彦・矢内 勇生. 2013. 『Stataによる計量政治学』オーム社.
分析はStataを用いて行われているが、著者ウェブサイト
(<http://yukiyanai.github.io/jp/quant-methods-stata/>) から対応するRコードを
入手可能
- Gandrudm, C. 2015. *Reproducible research with R and RStudio, Second edition*. Boca
Raton, FL: CRC Press.
- Grolemond, G. 2014. *Hands-on programming with R*. Sebastopol: O'Reilly (邦訳: 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)
- 飯田 健. 2013. 『計量政治分析』共立出版.
- 石田 基広・市川 太祐・瓜生 真也・湯谷 啓明. 2017. 『Rによるスクレイピング入門』C&R
研究所.
- Lander, JP. 2015[2014]. 『みんなのR: データ分析と統計解析の新しい教科書』高柳 慎一
ほか訳, マイナビ.
- Morgan III, JE. 2015. *Political analysis using R*. New York: Springer.
- R サポートズほか. 2017. 『パーフェクトR』技術評論社.
- 竹澤 邦夫. 2012. 『シミュレーションで理解する回帰分析』共立出版.
- Wickham, H. 2015. *Advanced R*. New York: Taylor & Francis (邦訳: 石田 基広ほか訳.
2016. 『R言語徹底解剖』共立出版)
- Wickham, H. & G. Grolemond. 2016. *R for data science: Import, tidy, transform,
visualize, and model data*. Sebastopol: O'Reilly Media.

副読本 (Rの解説を含まないもの・購入推奨)

- 伊藤 公一朗. 2017. 『データ分析の力: 因果関係に迫る思考法』光文社.
- 北村 行伸. 2009. 『ミクロ計量経済学入門』日本評論社.
- 鹿野 繁樹. 2015. 『新しい計量経済学: データで因果関係に迫る』日本評論社.
- 森田 果. 2014. 『実証分析入門: データから「因果関係」を読み解く作法』日本評論社.
- 末石 直也. 2015. 『計量経済学: ミクロデータ分析へのいざない』日本評論社.
- 田中 隆一. 2015. 『計量経済学の第一歩: 実証分析のススメ』有斐閣.
- 山本 勲. 2015. 『実証分析のための計量経済学: 正しい手法と結果の読み方』中央経済社.

ウェブサイト

- 青木 繁伸. 2009. 『Rによる統計解析』オーム社, サポートページ
<http://aoki2.si.gunma-u.ac.jp/R/>
- Coding Edge 「実践! Rで学ぶ統計解析の基礎」
<http://www.atmarkit.co.jp/fcoding/index/stat.html>
- Galili, Tal. “R tutorials”
<https://www.r-bloggers.com/how-to-learn-r-2/#h.y5b98o9o2h1r>
- R tips <http://cse.naro.affrc.go.jp/takezawa/r-tips/r.html>
- Wickham, H. & G. Grolemond. 2016. *R for Data Science: Import, Tidy, Transform,
Visualize, and Model Data*. サポートページ
<http://r4ds.had.co.nz>

成績評価の方法 / Evaluation

中間課題 2 回 (40%) と最終課題 (60%) による。課題の詳細は講義時に指示するが、いずれも R のコードと出力 (解析・可視化) 結果の報告・解釈を求める。当然ながら、正当な理由なく締切を過ぎた提出物は一切受け付けない。

関連科目 / Related course

計量経済学, 統計学, ミクロ経済学 I, ミクロ経済学 II, ゲーム分析, 金融論

リンク先ホームページアドレス / URL of syllabus or other information

- 講義資料ウェブページ: http://cfes-project.eco.u-toyama.ac.jp/education/education_2017/r_2017/
- R ウェブサイト: <https://www.r-project.org>
- R Studio ウェブサイト: <https://www.rstudio.com>