

統計学 2

高知工科大学 経済・マネジメント学群

2022 年度 1Q

開講日時：月曜・木曜 2 限

教室：A206

オフィスアワー：火曜 5 限

オフィスアワー以外の研究室訪問は**要予約**

担当：矢内 勇生（やない ゆうき）

研究室：A625

Email: yanai.yuki@kochi-tech.ac.jp

Website: <https://yukiyanai.github.io>

講義の概要と目的

現代社会にあふれるデータを適切に理解・処理する能力（統計リテラシー）を身につけ、社会を批判的に観察できる目を養うことがこの授業の目的である。統計のウソにだまされないだけでなく、自らも意図せざるウソをつかず、統計を適切に利用・提示できるようになることを目標とする。

履修要件

「統計学 1」を履修済みであり、四則演算、平方根、絶対値、冪乗（べきじょう）の意味がわかること。

関連科目

この科目は、「計量経済学」、「計量経済学応用」、「情報科学 3」の**前提科目**になっている。よって、「統計学 2」の単位を取得しないと今後の経済系の科目履修に支障が出る。そのことに十分留意して授業を受けること。「統計学 2」と上に挙げた 3 科目はすべて数学の教員免許取得のための選択科目になっているので、数学の教員免許取得を目指している場合は特に注意されたい。

授業の方法

この授業は、情報演習室（コンピュータ教室）で講義とコンピュータ実習を織り交ぜて行う。講義と実習の時間配分は内容に応じて変わる。**実習科目なので、対面授業の録画ならびにオンライン配信は行わない。**

実習では情報演習室に設置されているコンピュータを使うことができるが、自分のパソコンを持ち込んでもよい。自分のパソコンを持参する場合は、R（バージョン 4.1.2 を推奨¹⁾）と RStudio（バージョン 2022.02 以降を推奨）をあらかじめインストールしておくこと。また、電源コンセントの数に限りがある（数人分しかない）ので、授業前に十分に充電してくること。情報演習室のディスプレイは HDMI ポートが空いているので、HDMI 接続用のケーブルを持参すれば、自分のパソコンの画面と併せて 2 画面で授業を受けることができる。

成績評価

成績は以下の要素によって構成される（Plan A）。

- 授業への参加 [単なる出席は参加ではない]（最終成績の 10%）
- 課題の提出状況と完成度（45%）
- 期末試験（45%）
 - 期末試験の詳細については、授業中に案内する。

1) バージョン 4.0.0 以上がインストールされているなら無理にアップデートしなくてよい。講義開始時点での最新版は 4.1.3（2022 年 3 月 10 日公開）だが、情報演習室には 4.1.2 がインストールされているので、それに合わせて講義・実習を行う。4.0.4 は日本語が正しく表示されないので避けるべき。4.1.2 の macOS 用 (R-4.1.2.pkg) は <https://cran.r-project.org/bin/macosx/base/> で、Windows 用 (R-4.1.2-win.exe) は <https://cloud.r-project.org/bin/windows/base/old/4.1.2/> で入手できる。

これに加え、課題を提出したくないという学生のために期末試験の成績のみで成績を決める Plan B を用意する（100 点満点の期末試験の場合、60 点未満の者は F、60 点以上 70 点未満の者は C、70 点以上の者は B、80 点以上かつ順位が上位 25% 以内の者は A、90 点以上かつ順位が上位 5% 以内の者は AA）。すべての受講生について Plan A と Plan B の両方で点数を付け、良いほうの点数を採用する。ただし、統計分析手法を身につけ、卒業論文でデータ分析を行いたいという学生には、課題にも真剣に取り組むことを強く推奨する。

最終成績の目安は以下のとおり。

- AA** A の条件に加え、この授業で扱う統計学の定理（例：中心極限定理）や推定法（例：区間推定）の性質をシミュレーションによって確かめる R コードが書ける。[ただし、履修者の 5% 以内を上限とする]
- A** B の条件に加え、データを用いて統計的推定と検定を実際に行い、母集団について推論することができる。[ただし、AA と合わせて履修者の 25% 以内を上限とする]
- B** C の条件に加え、母集団と標本の違い、統計的推定ならびに統計的検定を正しく理解している。
- C** コンピュータ上のファイルに記録されたデータを用いて、記述統計の計算やデータの可視化を行うことができる。
- F** C の条件を満たさない場合

教科書

以下の教科書を全員入手すること（3Q の「計量経済学」でも引き続き使用する）。永国寺キャンパスの生協やインターネット書店などで購入することができる。最新版は第 1 版第 4 刷（2021 年 8 月 10 日発行）である。

- ・浅野正彦, 矢内勇生. 2018. 『R による計量政治学』オーム社.

また、R に関する副読本として、

- ・宋財法, 矢内勇生. 2022. 『私たちの R：ベストプラクティスの探求』(web book; 無料)

を利用する。

参考書

購入する必要はないが、授業内容の理解を助けると思われる本を以下に挙げる。教科書に加えて購入するなら、統計学については大屋（2020）を、データ可視化については Healy（2019）[の訳書] を特に勧める。Healy（2019）の原著草稿版はウェブで無料で読める。

- ・阿部真人. 2021. 『データ分析に必須の知識・考え方 統計学入門 仮説検定から統計モデリングまで重要トピックを完全網羅』ソシム.
- ・Freeman, Michael, and Joel Ross. 2019. *Programming Skills for Data Science: Start Writing Code to Wrangle, Analyze, and Visualize Data with R*. Princeton UP. (木村隆介 訳. 2021. 『データサイエンスのための R プログラミングスキル』共立出版.)
- ・Lander, Jared P. 2017. *R for Everyone: Advanced Analytics and Graphics*, Second Edition. Addison-Wesley. (高柳慎一ほか訳. 2018. 『みんなの R：データ分析と統計解析の新しい教科書 第 2 版』マイナビ.)
- ・Grolemund, Garret. 2014. *Hands-On Programming with R*. 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.)
- ・Healy, Kieran. 2019. *Data Visualization*. Princeton UP. (瓜生真也ほか訳. 2021. 『データ分析のためのデータ可視化入門』講談社.)
- ・石田基広. 2014. 『R 言語逆引きハンドブック 改訂 2 版』C&R 研究所.
- ・Kirk, Andy. 2019. *Data Visualisation: A Handbook for Data Driven Design*. SAGE. (黒川利明 訳. 2021.

『データビジュアライゼーション：データ駆動型デザインガイド』朝倉書店.)

- ・ 北川源四郎, 竹村彰通 (編) 2021. 『教養としてのデータサイエンス』講談社.
- ・ 小島寛之. 2006. 『完全独習統計学入門』ダイヤモンド社.
- ・ 久保川達也, 国友直人. 2016. 『統計学』東京大学出版会.
- ・ 大屋幸輔. 2020. 『コア・テキスト 統計学 第3版』新世社.
- ・ 松村優哉, 湯谷啓明, 紀ノ定保札, 前田和寛. 2021. 『R ユーザのための RStudio[実践] 入門: tidyverse によるモダンな分析フローの世界 [改訂 2 版]』技術評論社.

コンピュータの利用：R、RStudio、R Markdown

この授業では、オープンソースの統計処理言語である R の使い方を学習し、それを用いてデータの収集、管理、分析を行う。また、R を使うための統合開発環境 (IDE) として、RStudio を利用する。R 以外の分析ソフト (例: Stata, SPSS) や他の言語 (例: Python, Julia) の説明はしない。R と RStudio はどちらも無料であり、各自のコンピュータ (Linux, macOS, Windows) にインストールすることができる。

自分のパソコンにインストールする場合は、以下の資料を参照されたい。

- ・ Linux (Ubuntu): [PDF \(4.6MB\)](#)
- ・ macOS: [PDF \(4.9MB\)](#)
- ・ Windows: [PDF \(6.2MB\)](#)

Linux (Ubuntu) 編、macOS 編、Windows 編があるので、自分のパソコンの OS に合った資料を読むこと。Ubuntu 以外の Linux ユーザは、Ubuntu 編を参考にすれば自力でできると思われる。

課題や期末レポートの作成には必ず R Markdown を用いること。R Markdown については授業で使い方を説明するが、予習したい者は

- ・ 高橋康介. 2015. 「[連載：R Markdown で楽々レポートづくり](#)」技術評論社
- ・ 土井翔平. 2019. 「[R Markdown: pdf 編](#)」

などを参照。

授業のウェブサイト

実習で利用する R コード (R 言語によるプログラミング例) については、以下のウェブ資料で解説する。

<https://yukiyanai.github.io/stat2/>

講義で使用したスライドやその他の資料は KUTLMS で配布する。

Slack

授業時間外のコミュニケーションツールとして、Slack を使う。この授業の Slack ワークスペースは「KUT 統計学 2 (2022) <https://kut-stat2-2022.slack.com>」である。Slack の基本的な使い方については、

- ・ [Slack 初心者のためのクイックスタートガイド](#)
- ・ [【完全初心者向け！】コミュニケーションツール Slack \(スラック\) の使い方](#)
- ・ [初心者が Slack を使いこなすために知っておきたい 6 つのポイント](#)

などを参照されたい。

授業の内容についての質問は、Slack の適切なチャンネルに投稿すること。受講生には、質問するだけでなく、他の受講生の質問に積極的に回答することを期待する。回答は完全なものでなくてもかまわない。また、質問した後に自ら答えを見つけたときは、その答えを投稿し、他の受講生と共有することを求める。一定の時間をおいても回答がない場合や、R 言語に関する技術的な質問 (特定のコマンドを知っているかどうかだけが問題になる

ような場合) については、講義担当者が回答する。

Slack における質問、回答、議論は、授業への貢献とみなし、内容に応じて参加点を加算する。授業に無関係の内容や議論を妨害するような投稿でない限り、減点はしない。

授業に関する連絡は Slack を通じて行うので、議論に参加する気がなくても必ず Slack に参加し、投稿内容を定期的に確認すること。

以下のリンクをクリックすれば、この授業のワークスペースに登録できる。

<https://join.slack.com/t/kut-stat2-2022/signup>

ただし、登録には KUT のメールアドレス (@ugs.kochi-tech.ac.jp) が必要である。

メンバー登録するときに入力する名前（本名ならびに表示名）は匿名でもかまわないが、担当教員はメールアドレスから個人を特定できるので注意されたい。

SA（学生アシスタント）

この授業には実習を補助してくれる SA（学生アシスタント）がいるので、実習でわからないことがあれば SA にも質問してほしい。ただし、以下の注意を守ること。

- SA への質問は、コンピュータの使い方に関するものに限定する。統計学の内容については必ず教員に質問すること。
- 授業時間外に SA に対して授業に関する質問をすることは**禁止**する。SA が給料をもらえるのは授業時間中だけであり、授業時間外に質問に答える義務はない。
- SA には礼節をもって接すること。SA に対する暴言や暴力などは授業妨害であり、不正行為として扱う。

SA は過去にこの授業の単位を優秀な成績で取得した先輩たちである。今年度の成績優秀者には、来年度以降の SA をお願いすることがある。

授業計画

授業計画は以下の通りである。毎回の**予習課題を読んでから授業に参加**すること。ここに挙げたもの以外の参考文献がある場合は KUTLMS で紹介する。

Topic 1. イントロダクション（第 1 回）[4/11 月]

まず、授業の概要、進め方、成績評価の方法などについて確認する。その後、統計学を学ぶ意義について考える。

予習・復習 必要なし

参考 読書案内 (p.6) に挙げられた本を 1 冊以上読むことを推奨する

Topic 2. R の基本操作（第 2 回）[4/14 木]

R を使って統計分析を行うための基礎を身につける。R の操作法を学びながら、統計学 1 の復習をする。

準備 R のインストール（自分のパソコンを使う場合のみ）

- 担当教員のウェブサイト**にある資料（上に書いたもの）を参考にして、自分のパソコンに R と RStudio をインストールする。
- インストールがうまくいかない場合は、情報演習室 (A206, A204) のパソコン（R と RStudio がインストール済み）を使うこと。

予習・復習 教科書 第 4 章

参考 高橋 (2018) 第 2 章

参考 Healy (2019 [2021]) 第 2 章

Topic 3. 記述統計とデータの可視化 (第 3 回 [4/18 月], 第 4 回 [4/21 木])

記述統計の復習をする。さらに、データを可視化する基本的な方法と可視化の際に注意すべき事項を学ぶ。

予習・復習 教科書 第 5–6 章

予習・復習 Healy (2019 [2021]) 第 1 章

参考 永田ゆかり. 2020. 『データ視覚化のデザイン』SB クリエイティブ.

参考 Excel のダメなグラフでウソをつく法 [<http://id.fnshr.info/2016/03/28/excel-bad-charts/>]

参考 このグラフがひどい！ 2018 [<http://hidograph.com/2018/winners/>]

参考 なぜ円グラフを安易に使ってはいけないのか？ [<https://jikitourai.net/dont-use-piechart>]

Topic 4. R Markdown によるレポート作成 (第 5 回 [4/25 月])

R Markdown を利用して、RStudio でレポートを作成する方法を学ぶ。

予習・復習 高橋 (2018) 第 3–4 章

参考 高橋康介. 2015. 「[連載：R Markdown で楽々レポートづくり](#)」技術評論社.

Topic 5. ggplot2 入門 (第 6 回 [4/28 木], 第 7 回 [動画 (オンデマンド)])

R のパッケージである ggplot2 を使ってグラフを作る方法を学習する。

予習・復習 Healy (2019 [2021]) 第 3 章

復習 教科書 第 6 章

参考 Chang (2018)

Topic 6. シミュレーション (第 8 回 [5/12 木], 第 9 回 [5/16 月])

R で乱数を生成する方法と、それを利用してシミュレーションを行う方法を学ぶ。例として、シミュレーションによって中心極限定理の意味を理解する。

予習・復習 教科書 第 7–8 章

Topic 7. 統計的推定と仮説検定の基礎 (第 10 回 [5/19 木], 第 11 回 [5/23 月])

推測統計学の考え方に触れ、統計的推定と検定の手続きの概要を理解する。

予襲・復習 教科書 第 7–8 章

参考 小島『完全独習統計学入門』pp.90–106

Topic 8. 標本平均と母平均 (第 12 回 [5/26 木], 第 13 回 [5/30 月])

標本分布と標準誤差、点推定と区間推定について理解し、標本平均から母平均を推定する方法を学ぶ。

予習・復習 教科書 第 7–8 章

参考 小島 (2006) pp.110–148

Topic 9. t 分布と母平均の推定 (第 14 回 [6/2 木])

t 分布について理解し、標本から母集団の平均 (母平均) を推定する方法を学ぶ。

予習・復習 教科書 第 7 章

参考 小島 (2006) pp.180–195

Topic 10. 2 つの平均値を比較する (第 15 回 [6/6 月])

2 つのグループの間で平均値を比較し、統計的に差があるといえるかどうか確かめる方法を習得する。

予習・復習 教科書 第 8 章

予習・復習 浅野・矢内. 2013. 『Stata による計量政治学』 第 8 章.

単位認定試験（期末試験） [6/9 木]

教科書と A4 用紙 1 枚のチートシートのみ持込可。チートシートについては授業で説明する。

読書案内

統計学に関する様々な話題に関する読み物を以下にいくつか挙げる。この中から何冊か読んでみると、統計を勉強する意義や統計の面白さがわかるだろう。

- ・ダレル・ハフ『統計でウソをつく法』（1968 年、講談社ブルーバックス）
- ・竹内薫『99.9%は仮説』（2006 年、光文社新書）
- ・アンソニー・ルーベン『統計的な？数字に騙されないための 10 の視点』（2019 年、すばる舎リンケージ）
- ・チャールズ・ウィーラン『統計学をまる裸にする：データはもう怖くない』（2014 年、日本経済新聞出版）
- ・マイケル・ブラストランド、デイヴィッド・シュビーゲルハルター『もうダメかも：死ぬ確率の統計学』（2020 年、みすず書房）
- ・デイヴィッド・サルツブルグ『「誤差」「大間違い」「ウソ」を見分ける統計学』（2021 年、共立出版）
- ・ジェフリー・S・ローゼンタール『それはあくまで偶然です：運と迷信の統計学』（2021 年、早川書房）
- ・カイザー・ファング『ヤバい統計学』（2011 年、阪急コミュニケーションズ）
- ・オリヴィエ・レイ『統計の歴史』（2020 年、原書房）
- ・ウィリアム・パウンドストーン『世界を支配するベイズの定理：スパムメールの仕分けから人類の終焉までを予測する究極の方程式』（2020 年、青土社）
- ・バート・K・ホランド『確率・統計で世界を読む』（2004 年、白揚社）
- ・西内啓『統計学が最強の学問である』（2013 年、ダイヤモンド社）
- ・結城浩『数学ガールの秘密ノート / やさしい統計』（2016 年、SB クリエイティブ）
- ・ネイト・シルバー『シグナル&ノイズ：天才データアナリストの予測学』（2013 年、日経 BP）
- ・C.R. ラオ『統計学とは何か：偶然を生かす』（2010 年、ちくま学芸文庫）
- ・マイケル・ルイス『マネー・ボール』（2013 年、ハヤカワ・ノンフィクション文庫）
- ・鳥越規央『統計学が見つけた野球の真理：最先端のセイバーメトリクスが明らかにしたもの』（2022 年、講談社ブルーバックス）
- ・丸山健夫『ナイチンゲールは統計学者だった！：統計の人物と歴史の物語』（2008 年、日科技連出版社）
- ・石黒真木夫『統計学をめぐる散歩道：ツキは続く？ 続かない？』（2020 年、岩波ジュニア新書）