

統計学 1

高知工科大学 経済・マネジメント学群 2024 (令和 6) 年度 1Q

開講日時：火曜・金曜 3 限

担当：矢内 勇生 (やない ゆうき)

教室：A206

研究室：A625

オフィスアワー：火曜 5 限 (16:50-18:20)

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

オフィスアワー以外の研究室訪問は要予約

Website: <https://yukiyamai.github.io/jp/>

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

経済学・マネジメント学のための数学 I および II (「線形代数学 1, 2」ならびに「微分積分学 1, 2」または「数学 1, 2, 4」でも可) の単位を取得済みであること。

関連科目

この科目は、「統計学 2」、「計量経済学」、「計量経済学応用」の前提科目になっている。よって、「統計学 1」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。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。

在学中に統計分析・実証分析の手法を身に付けたいなら、「統計学 1」→「アンケート調査法」→(「確率・統計」→)「統計学 2」→「実験デザイン」→「計量経済学」→「プログラミング」→「計量経済学応用」→「情報科学 3」を 2 年かけて履修することを推奨する。ただし、年度によっては科目の提供順が前後する場合があります。

授業の方法

この授業は、ハイブリッド形式で実施する。講義 (の大部分) についてはオンライン (KUTLMS) で動画を提供する。対面授業は、情報演習室で実習を織り交ぜながら行う。受講生は、情報演習室に備え付けのコンピュータを使って実習を行う。

対面授業は受講生を A 班と B 班の 2 つに分けて行う。班分けは KUTLMS にアップロードする。表 1 に示すとおり、A 班と B 班では対面授業に出席すべき日程が異なるので注意されたい。自分の班の日程に出席できないときは、他の班の実習に参加してよい。また、希望者は両日程に参加してもよい。ただし、座席が不足する場合には該当班の学生を優先するので、**できるだけ指定された日程で参加してほしい** (各班の人数がクラス全体の約半分になるようにしてくれれば、学生同士で相談して班分けを変更してくれても問題ない)。

実習では情報演習室に設置されているコンピュータを利用する。自分のパソコンを持ち込んでもよいが、**個人所有のコンピュータで発生するエラーには対処しない¹⁾**。自力でエラーを解決できる者以外は情報演習室に設置されているコンピュータを利用すること²⁾。**対面授業が始まる 4 月 12 日までに、コンピュータへのログイン方**

1) 時間がないので、「統計学 2」以降では個人所有のパソコンの利用も認める予定 (受講人数による)。

2) 自分のパソコンを持参する場合は、R (バージョン 4.3.3 を推奨) と RStudio (バージョン 2023.12.1 以降を推奨) をあらかじめインストールしておくこと。講義開始時点での最新版は 4.3.3 (2024 年 2 月 29 日公開) だが、情報演習室には 4.3.2 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。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おくこと。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。

法を必ず確認しておくこと。

成績評価

成績は以下の要素によって構成される (Plan A)。

- 課題の提出状況と完成度 (40%)
- 期末試験 (60%)
 - 期末試験の詳細については、授業中に案内する。

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、様々な確率分布の特徴を把握するための R コードを書くことができる。[ただし、履修者の 5% 以内を上限とする]
- A** B の条件に加え、確率モデルを使ってデータから情報を引き出すことができる。[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、離散型分布と連続型分布の基本的な性質を理解してる。
- 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「統計学 2」、3 年生担当の「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 6 刷 (2023 年 4 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. 2024. 『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される本を以下に挙げる。教科書に加えて購入するならば、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.)
- Golemund, Garret. 2014. *Hands-On Programming with R*. O'Reilly. (大橋真也 監訳. 2015. 『RStudio で始める 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.
- 岩沢宏和. 2016. 『分布からはじめる確率・統計入門：実用のための直感的アプローチ』東京図書.
- 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.
- 溝渕健一, 谷崎久志. 2021. 『統計学』ミネルヴァ書房.
- 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、Quarto、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。

課題の作成には必ず Quarto (または R Markdown) を用いること。 Quarto については授業で使い方を説明するが、予習したい者は『私たちの R』の「[Quarto 入門](#)」を参照されたい。

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/stat1/>

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 1 (2024) <https://kut-stat1-2024.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。

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-stat1-2024/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表示名は本名にすること (本名でないと課題の返却時に困るので)。本名の代わりに匿名を利用したい者は担当教員に DM で連絡すること。

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担当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。ただし、授業の進捗状況に応じて変更する可能性がある。変更する際はこの講義要綱を更新し、授業中に案内する。各回の内容に応じて、予習、復習で読むべき教科書の範囲を示すので、必ず予習・復習すること。**予習課題は全員読んでいるという前提で授業を進める**。講義の前に**予習課題を読み**、該当するトピックの**講義動画を視聴してから対面授業に参加**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。

表 1 授業スケジュール

授業 No.	対象トピック	形式	A 班対面授業	B 班対面授業	備考
1	1	動画			
2	2	対面	4 月 12 日 (金)		2 班合同
3	2	対面	4 月 16 日 (火)	4 月 19 日 (金)	
4	3	動画			
5	3	動画			
6	3	対面	4 月 23 日 (火)	4 月 26 日 (金)	
7	4	動画			
8	4	対面	4 月 30 日 (火)	5 月 7 日 (火)	
9	5	動画			
10	5	対面	5 月 10 日 (金)	5 月 14 日 (火)	
11	6	動画			
12	6	対面	5 月 17 日 (金)	5 月 21 日 (火)	
13	7	動画			
14	7	対面	5 月 24 日 (金)	5 月 28 日 (火)	
15	–	期末試験	5 月 31 日 (金)		試験教室に注意
16	8	対面	6 月 4 日 (火)		2 班合同

講義動画は [KUTLMS](#) にアップロードする。

Topic 1. イントロダクション (第 1 回)

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 (このシラバスの末尾) 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 2–3 回)

R を使って統計分析を行うための基礎を身につける。R の基本操作と、RStudio の利用方法について学ぶ。

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

参考 宋・矢内「[私たちの R](#)」

Topic 3. データの種類と記述統計 (第 4–6 回)

統計分析で扱うデータの種類について理解し、R を用いてデータに含まれる情報を要約する方法を学ぶ。さらに、Quarto を用いてレポートを作成する方法を学ぶ。

予習・復習 教科書 第 5–6 章

予習・復習 Healy (2019 [2021]) 第 1 章

予習・復習 高橋 (2018) 第 3–4 章

参考 吉田寿夫. 1998. 『本当にわかりやすいすごく大切なことが書いてあるごく初歩の統計の本』北大路書房：序章

参考 上田雅夫, 後藤正幸. 2022. 『データサイエンス入門』有斐閣：第 2 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 アルベルト・カイロ. (数井真澄 訳) 2020. 『グラフのウソを見破る技術』ダイヤモンド社.

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>]

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 4. 離散分布 (第 7–8 回)

確率分布について学ぶ。特に、離散型の確率分布について学習し、確率分布の意味と確率分布を視覚的に捉える方法を理解する。

予習・復習 栗原伸一, 丸山敦史. 2017. 『統計学図鑑』オーム社：第 2 章

予習・復習 岩沢 (2016) 第 1–2 章

参考 大屋 (2020) 第 4 章

参考 浜田宏一. 2018. 『その問題、数理モデルが解決します』序章 – 第 2 章

Topic 5. 連続分布 (第 9–10 回)

連続型の確率分布について学ぶ。離散型と連続型の確率分布の違いを理解する。

予習・復習 岩沢 (2016) 第 3 章

予習・復習 溝淵・谷崎 (2021) 第 4 章

参考 大屋 (2020) 第 5 章

Topic 6. 確率変数と確率分布 (第 11–12 回)

確率変数について学び、確率分布がどのように利用されるか理解する。

予習・復習 阿部 (2021) 第 3 章

予習・復習 宮川公男. 2022. 『基本統計学 第 5 版』有斐閣：第 5 章 (pp.119–132)

参考 岩沢 (2016) 第 4 章

参考 大屋 (2020) 第 3 章

参考 浜田宏. 2020. 『その問題、やっぱり数理モデルが解決します』2–4 章

Topic 7. 分布の代表値 (第 13–14 回)

確率分布を要約する方法について学び、代表的な確率分布の特徴を理解するとともに、それらの特徴を生かすと確率分布をどのような場面で利用することができるかについて考える。

予習・復習 阿部 (2021) 第 3 章 (再読)

予習・復習 溝淵・谷崎 (2021) 第 2–3 章

参考 岩沢 (2016) 第 5 章

参考 久保川・国友 (2016) 2 章

単位認定試験（期末試験） [5/31（金）]

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。

Topic 8. 全体のまとめ（第15回）

復習 1Qの内容すべて

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・シュピーゲルハルター『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・シュピーゲルハルター『統計学の極意』（2024年、草思社）
- ・ 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）