

4th Apr 2022

スタートアップゼミ

研究の仕方

TeX の使い方, R の使い方

小林里瑛

本日の流れ



研究の骨子: 要素を埋められるか？

背景

課題

目的

困難

アイデア

方法論・実証

結果

結論

研究の骨子

背景

- 実社会における研究の意義
- 同じ問題に取り組んでいる研究群に対する研究の意義

課題

- 既往研究は何を明らかにしているのか
- 既往研究は何を明らかにしていないのか

目的

- 研究で何を新たに明らかにしようとしているのか or 何を新たに達成しようとしているのか

困難

- 明らかにする or 達成するためにどんな困難があるのか

アイデア

- 困難を解決するためにどんなアイデアがあるのか
- そのアイデアには新規性があるか

方法論・実証

- アイデアを生かし困難に立ち向かうための実験のデザインは何か
- 仮説を検証するための実験のデザインは何か

結果

- どんなデータが得られた（得られそう）か
- データから言える（言えそうな）ことは何か

結論

- 目的を明らかにすることで不足点にどのような貢献ができるか

研究の骨子



背景

- 実社会における研究の意義
- 同じ問題に取り組んでいる研究群に対する研究の意義

課題

- 既往研究は何を明らかにしているのか
- 既往研究は何を明らかにしていないのか

目的

- 研究で何を新たに明らかにしようとしているのか or 何を新たに達成しようとしているのか

行ったり来たりしながら進めます

方法論・実証

- アイデアを生かし困難に立ち向かうための実験のデザインは何か
- 仮説を検証するための実験のデザインは何か

結果

- どんなデータが得られた（得られそう）か
- データから言える（言えそうな）ことは何か

結論

- 目的を明らかにすることで不足点にどのような貢献ができるか

参考 : Nature の推奨するアブストラクトの書き方

1. 背景
2. 詳細な背景
3. 対象とする問題
4. 結果の要約
5. 結果の詳細
6. 結果の一般化
7. 展望

1～2文で、基本的な紹介を、全ての分野の科学者が理解できるように記載する。

2～3文で、研究のより詳細な背景を、関連分野の科学者が理解できるように記載する。

1文で、この論文の研究が対象としている一般的な問題を明確に述べる。

1文で、主要な結果を要約する。

2～3文で、今回明らかになった主要な結果を、従来はどのように考えられていたかということと直接比較しながら説明する。

1～2文で、結果をより一般的な内容に落とし込む。

2～3文で、全ての分野の科学者が容易に理解できるような、より広範な展望を記載することがある（アブストラクトにこうした展望が含まれることで、その論文の影響力が大きく増すと編集者が判断する場合）。この場合アブストラクトは、最大で300語程度となる。

During cell division, mitotic spindles are assembled by microtubule-based motor proteins^{1,2}. The bipolar organization of spindles is essential for proper segregation of chromosomes, and requires plus-end-directed homotetrameric motor proteins of the widely conserved kinesin-5 (BimC) family³. Hypotheses for bipolar spindle formation include the 'push-pull mitotic muscle' model, in which kinesin-5 and opposing motor proteins act between overlapping microtubules^{2,4,5}. However, the precise roles of kinesin-5 during this process are unknown. Here we show that the vertebrate kinesin-5 Eg5 drives the sliding of microtubules depending on their relative orientation. We found in controlled *in vitro* assays that Eg5 has the remarkable capability of simultaneously moving at $\approx 20 \text{ nm s}^{-1}$ towards the plus-ends of each of the two microtubules it crosslinks. For anti-parallel microtubules, this results in relative sliding at $\approx 40 \text{ nm s}^{-1}$, comparable to spindle pole separation rates *in vivo*⁶. Furthermore, we found that Eg5 can tether microtubule plus-ends, suggesting an additional microtubule-binding mode for Eg5. Our results demonstrate how members of the kinesin-5 family are likely to function in mitosis, pushing apart interpolar microtubules as well as recruiting microtubules into bundles that are subsequently polarized by relative sliding. We anticipate our assay to be a starting point for more sophisticated *in vitro* models of mitotic spindles. For example, the individual and combined action of multiple mitotic motors could be tested, including minus-end-directed motors opposing Eg5 motility. Furthermore, Eg5 inhibition is a major target of anti-cancer drug development, and a well-defined and quantitative assay for motor function will be relevant for such developments.

引用 : <https://storage.googleapis.com/natureasia-assets/pdf/ja-jp/nature/authors/gta-2017-updated.pdf>

Texとは

- 論文や本を書くとき, 文章や図表を綺麗に組版処理してくれる言語でありシステム
 - 類似の機能を持つソフトとしてAdobe社のInDesignがある
- 画像以外の本文や書式設定を全てテキストファイルで記述
- ソースファイルをコンパイルしてPDFファイルを生成
- “テフ” もしくは “テック” と読む

L^AT_EX
“ラ” “テフ”

TeXとは

TeX を使うメリット

- 図表番号を勝手に振ってくれる
- 図表を入れた後に文章を加筆・削減しても図表がずれない
- 数式を綺麗に書くことができる
- 参考文献の管理が簡単

TeX の使い方

- オフラインで使う → LaTeX を DL + Atom
- オンラインで使う → Overleaf

ヘッダー・ページ番号

64

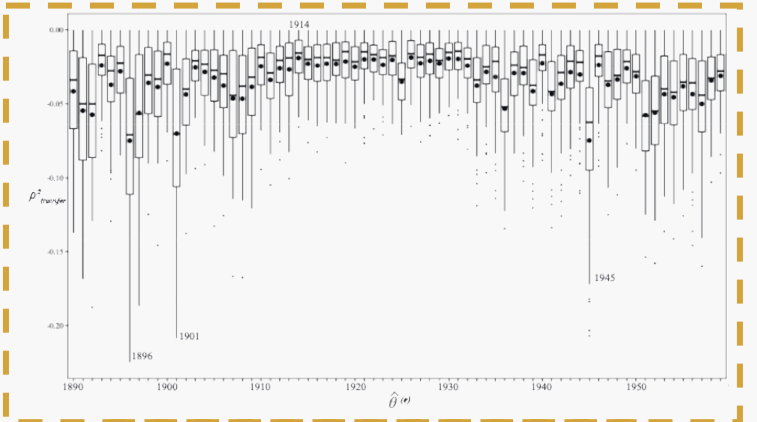
Chapter 4. Long-term Latent Class Cross Nested Land Choice Model

The housing parameter ratio β_H/β_C is the parameter that shows the most difference between the classes, although there are some periods of high variability. Between 1930 and 1940, the development of low-use land in the periphery increased, and land transactions became more active than in the former period. The class with a positive parameter ratio between 1930 and 1935 is considered to be the class that preferred the land ownership pattern that includes low-use land due to this development. The results show that, even if the background is unknown, we can extract the class that prefers land that includes low-use land.

小見出し

4.5.4 Validation

図



図表番号

FIGURE 4.8: transfer rho-square for each dataset $v = [1890, 1959]$

参考文献

To test temporal transferability, we validate estimated models. For the classification and definition of validation in the Behavior model, please refer to Parady et al. (2021). This subsection examines the external validation. However, the samples which compose the dataset are almost the same before and after a particular year. Therefore, the internal validation will be partly examined.

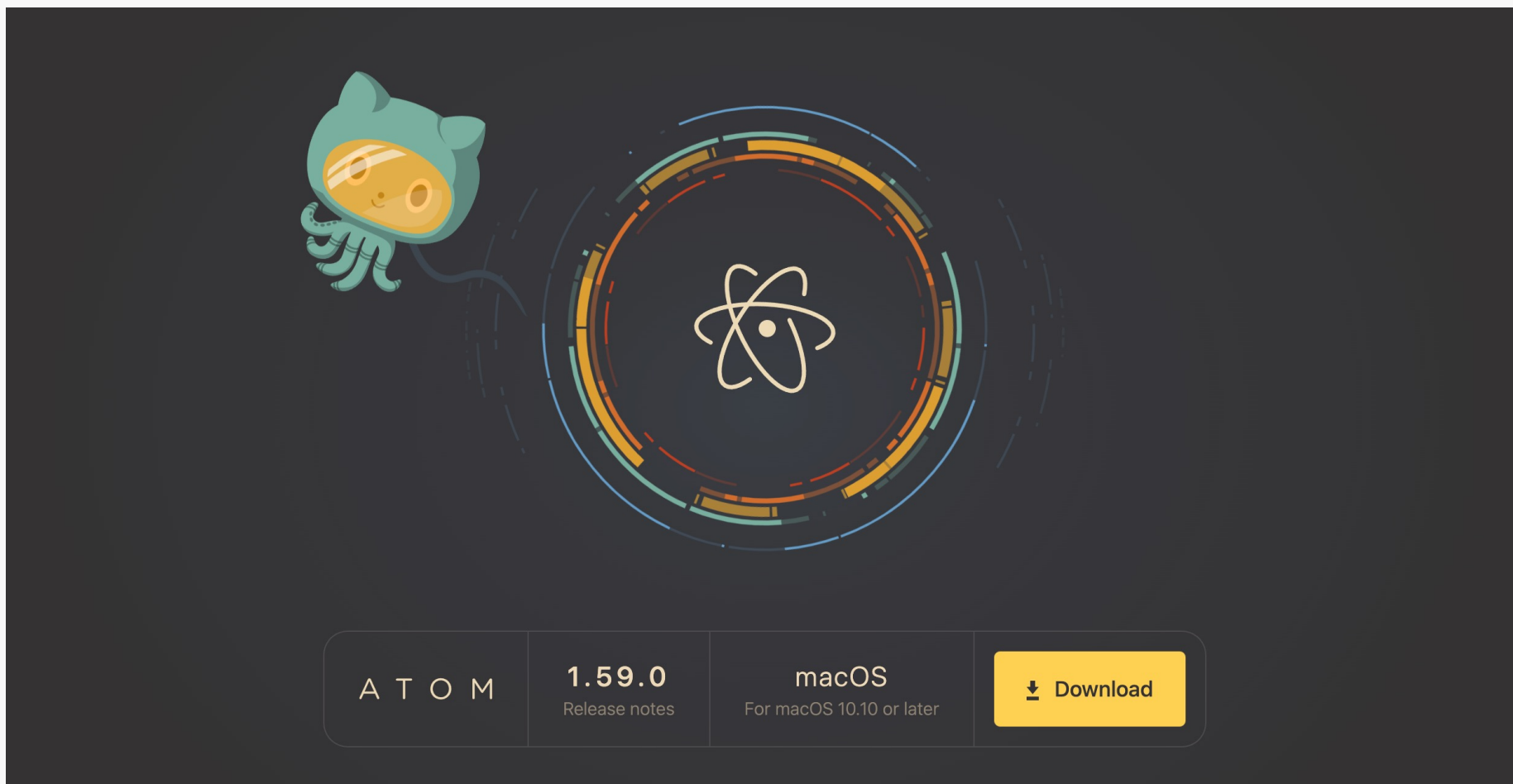
We employ transfer rho-square as performance measures. The transfer rho-square is absolute and easy to use as a benchmark. The definition of transfer rho-square is as follow:

数式と数式番号

$$\rho_{transfer}^{2,(e),v} = 1 - \frac{LL_v(\hat{\theta}^{(e)})}{LL_v(\hat{\theta}^{(v)})} \quad (4.22)$$

ローカルでLatex: エディタにAtom を使う場合 for mac

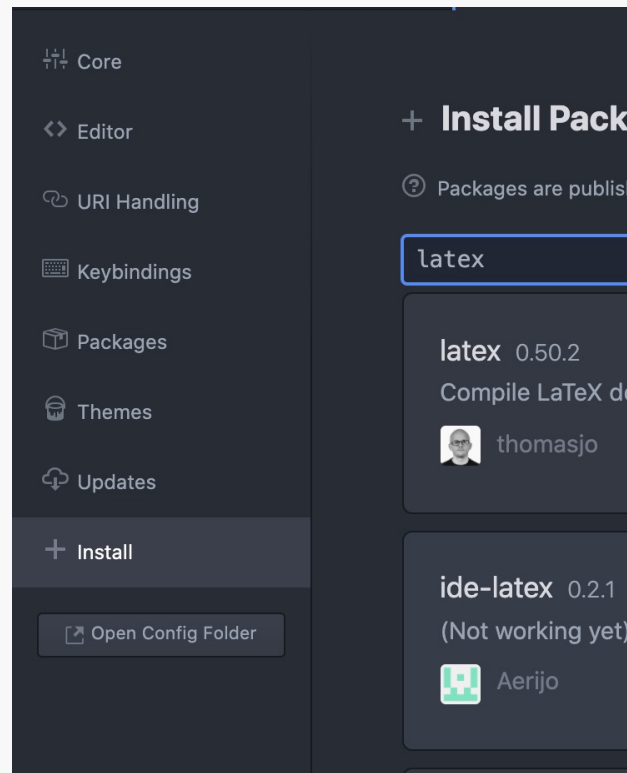
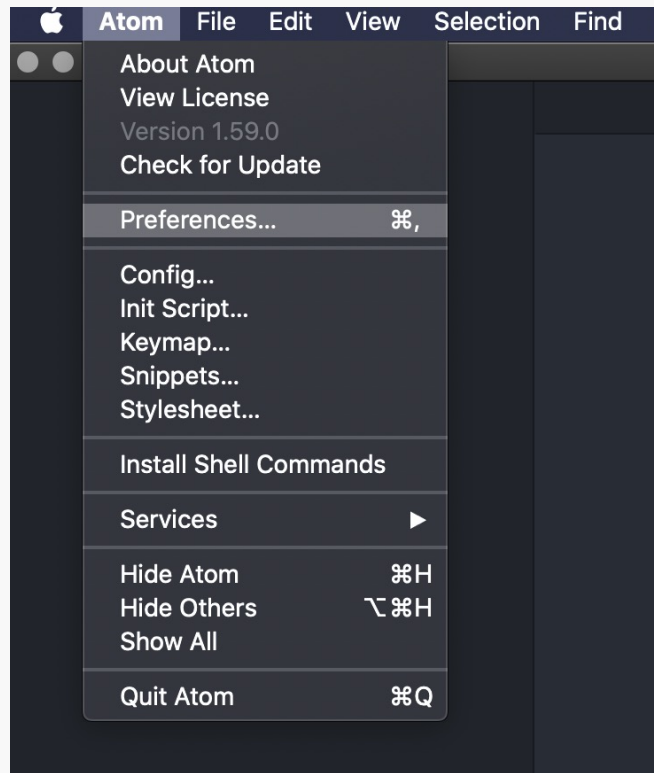
1. Atom を <https://atom.io/> からダウンロードする



ローカルでLatex: エディタにAtom を使う場合 for mac

2. Atom 上でパッケージをインストールする

- Atom を開き, Atom > Preference を選択
- 左メニューの一番下「Install」から次のパッケージをインストール



- latex
- language-latex
- pdf-view

ローカルでLatex: エディタにAtom を使う場合 for mac

3. Homebrew をインストールする

- ターミナルを開き, コマンドを入力し, Enter を押す

```
% /bin/bash -c "$(curl -fsSL  
https://raw.githubusercontent.com/Homebrew/install/HEAD/install.sh)"
```

- インストールが完了したら, 指示通りにパスを以下のコマンドにて通す

```
% echo 'eval "$(/opt/homebrew/bin/brew shellenv)'" >>  
/Users/bin/.zprofile eval "$(/opt/homebrew/bin/brew shellenv)"
```

- Cask を導入する

```
brew tap homebrew/cask
```

ローカルでLatex: エディタにAtom を使う場合 for mac

4. MacTeX をインストールする

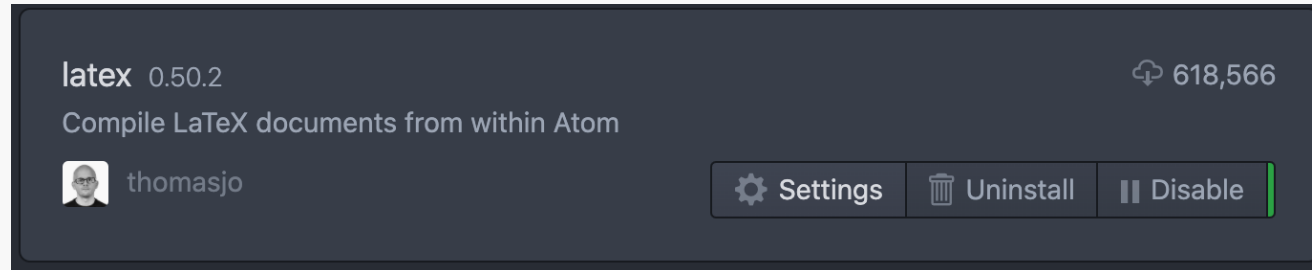
- ターミナルでコマンドを入力し, インストール完了まで待つ

```
brew install --cask mactex
```

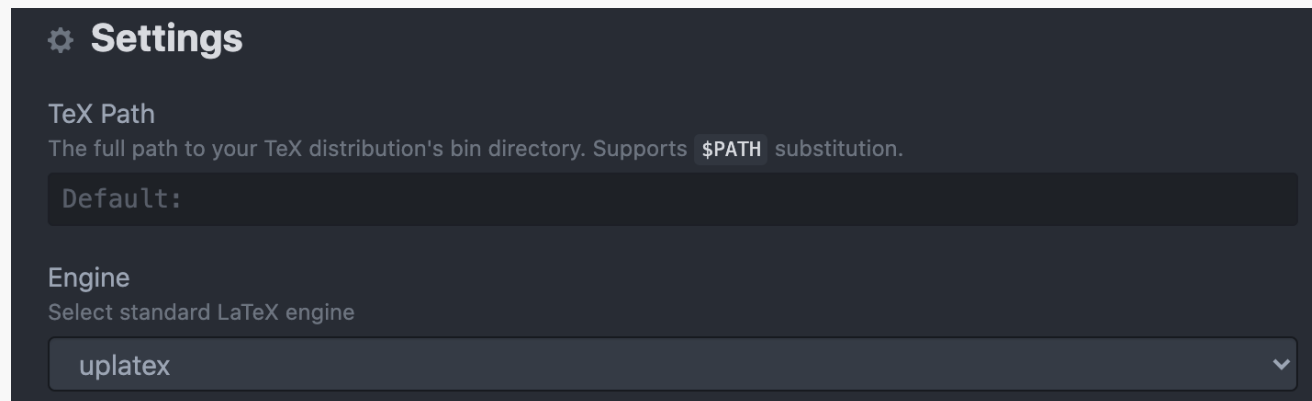
ローカルでLatex: エディタにAtom を使う場合 for mac

5. Atom 側で設定

- Atom > Preference > Packages > latex > Settings を選択



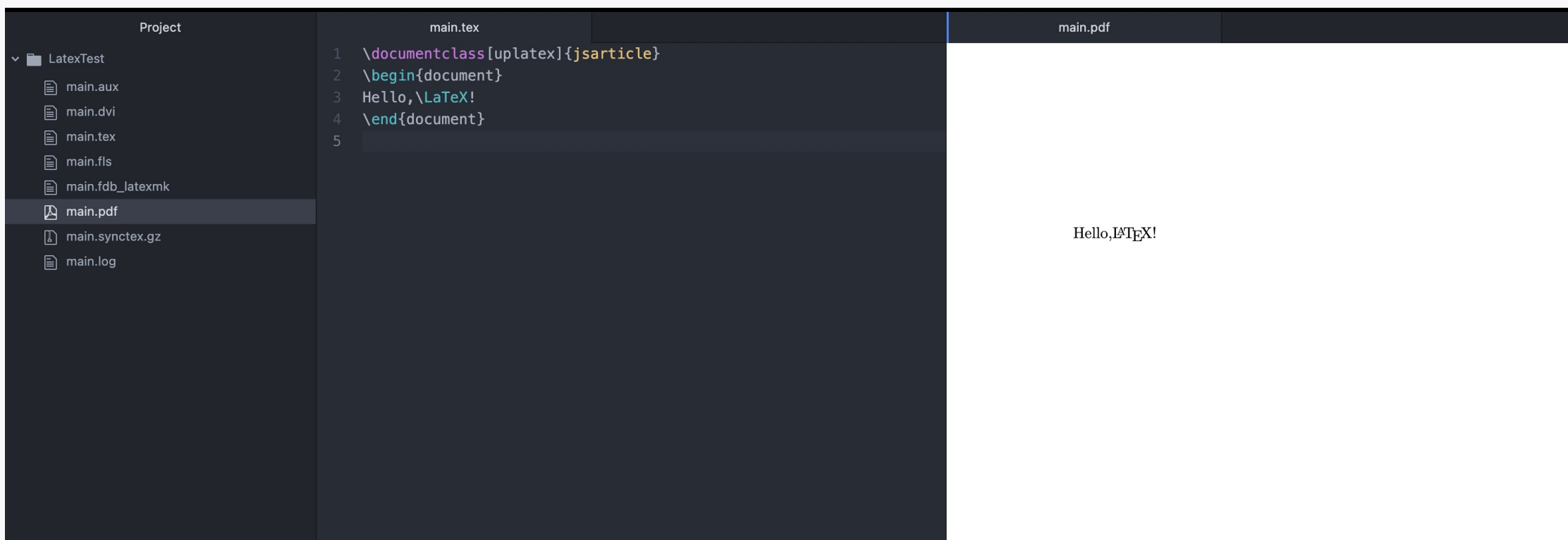
- Engine を “uplatex” に設定



ローカルでLatex: エディタにAtom を使う場合 for mac

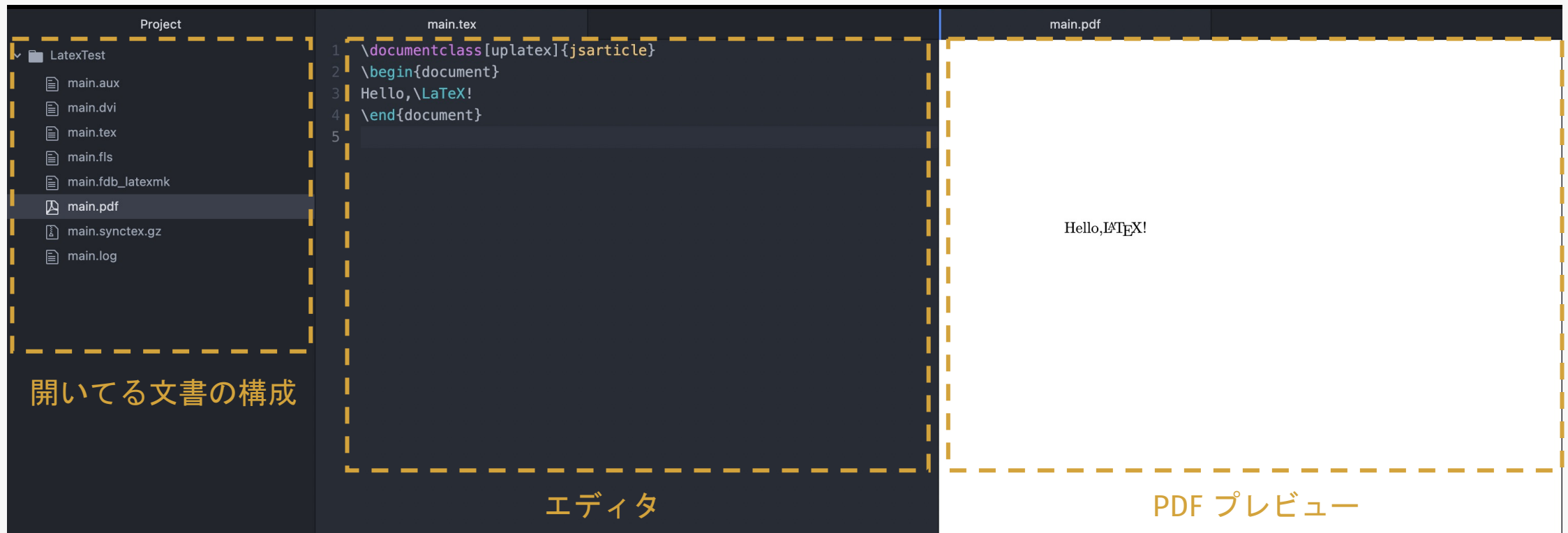
6. 動作確認

- "main.tex" を作成し、コンパイルするとPDF が生成される



ローカルでLatex: エディタにAtom を使う場合 for mac

エディタAtom の構成



オンラインでLatex: Overleaf の導入

- Overleaf を使うメリット

- ローカルでの環境構築が不要
- クラウド上に保存可能なので、どこでも執筆が可能
 - 逆に言えばオフラインでは使えない
- 招待することで共同作業が可能
- Git を使うことで変更履歴を管理可能
- テンプレートが多数配布されている
 - スライドを作ることもできます →
 - 数式の多い資料を作るときは便利かも



テンプレートの例 ▶
beamer : “metropolis”

Recursive Macroeconomics

Chapter1:Overview

Chapter.6:Search and Unemployment

Risa Kobayashi

2021/05/28

Recursive Macroeconomic ゼミ #1

オンラインでLatex: Overleaf の登録

登録する

Register using your email

I'd like emails about product offers and company news and events.

or



Register using Google



Register using ORCID

or

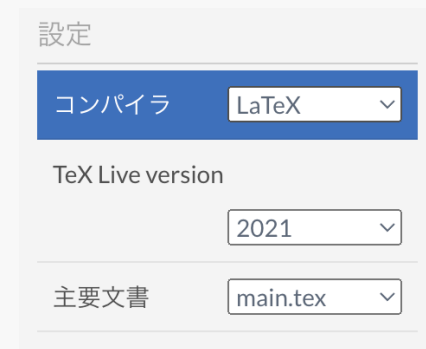
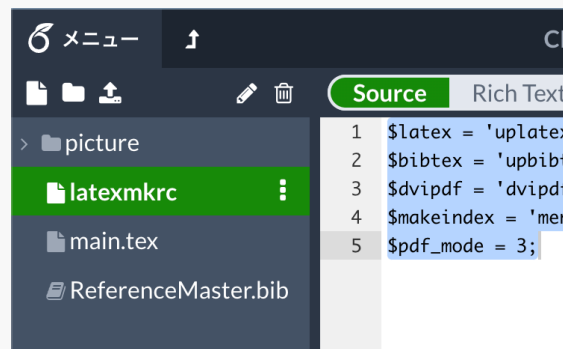
[Log in through your institution](#)

オンラインでLatex: Overleaf で日本語論文を書く

1. latexmkrc ファイルをmain.tex と同じ階層に作成
2. 以下の内容をファイル latexmkrc にコピーする

```
$latex = 'uplatex';  
$bibtex = 'upbibtex';  
$dvi2pdf = 'dvi2pdfmx %O -o %D %S';  
$makeindex = 'makeindex -U %O -o %D %S';  
$pdf_mode = 3;
```

3. コンパイラをLatexに変更する



オンラインでLatex: Overleaf の構成

Menu ↑ DoctoralThesis_校正済

Source Rich Text

Appendices
Chapters
Figures
Thesis_Additional
algorithm.sty
algorithmic.sty
cite.sty
jpncolor.tex
jsarticle.cls
latexmkrc
main.tex
MastersDoctoralThesi.
model2-names.bst
natbib.sty
ReferenceMaster.bib

```
1 %~~~~~  
2 % Masters/Doctoral Thesis  
3 % LaTeX Template  
4 % Version 2.5 (27/8/17)  
5 %  
6 % This template was downloaded from:  
7 % http://www.LaTeXTemplates.com  
8 %  
9 % Version 2.x major modifications by:  
10 % Vel (vel@latextemplates.com)  
11 %  
12 % This template is based on a template by:  
13 % Steve Gunn (http://users.ecs.soton.ac.uk/srg/softwaretools/document/templates/)  
14 % Sunil Patel (http://www.sunilpatel.co.uk/thesis-template/)  
15 %  
16 % Template license:  
17 % CC BY-NC-SA 3.0 (http://creativecommons.org/licenses/by-nc-sa/3.0/)  
18 %  
19 %~~~~~  
20 %  
21 %-----  
22 % PACKAGES AND OTHER DOCUMENT CONFIGURATIONS  
23 %-----  
24 %  
25 \documentclass[  
26 11pt, % The default document font size, options: 10pt, 11pt, 12pt  
27 %oneside, % Two side (alternating margins) for binding by default, uncommand to switch to one side  
28 english, % ngerman for German  
29 singlespacing, % Single line spacing, alternatives: onehalfspacing or doublespacing  
30 %draft, % Uncomment to enable draft mode (no pictures, no links, overfull hboxes indicated)  
31 %nolistspacing, % If the document is onehalfspacing or doublespacing, uncommand this to set spacing in lists to  
32 single  
33 %liststotoc, % Uncomment to add the list of figures/tables/etc to the table of contents  
34 %tactotoc, % Uncomment to add the main table of contents to the table of contents  
35 %parskip, % Uncomment to add space between paragraphs  
36 %nohyperref, % Uncomment to not load the hyperref package  
37 headsepline, % Uncomment to get a line under the header  
38 %chapterinonline, % Uncomment to place the chapter title next to the number on one line  
39 %consistentlayout, % Uncomment to change the layout of the declaration, abstract and acknowledgements pages to  
40 match the default layout  
41 ]{MastersDoctoralThesis} % The class file specifying the document structure  
42 %  
43 \usepackage[utf8]{inputenc} % Required for inputting international characters  
44 \usepackage[T1]{fontenc} % Output font encoding for international characters  
45 %  
46 \usepackage{mathpazo} % Use the Palatino font by default  
47 %  
48 \usepackage[backend=bibtex,style=authoryear,natbib=true]{biblatex} % Use the bibtex backend with the authoryear  
49 citation style (which resembles APA)
```

エディタ

Review Share Submit History Cha

Recompile

64 Chapter 4. Long-term Latent Class Cross Nested Land Choice Model

The housing parameter ratio β_H/β_C is the parameter that shows the most difference between the classes, although there are some periods of high variability. Between 1930 and 1940, the development of low-use land in the periphery increased, and land transactions became more active than in the former period. The class with a positive parameter ratio between 1930 and 1935 is considered to be the class that preferred the land ownership pattern that includes low-use land due to this development. The results show that, even if the background is unknown, we can extract the class that prefers land that includes low-use land.

4.5.4 Validation PDF プレビュー

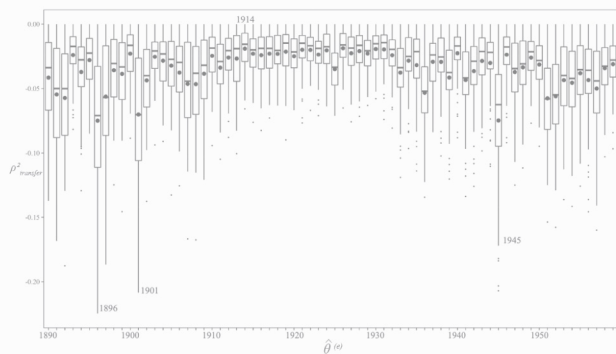


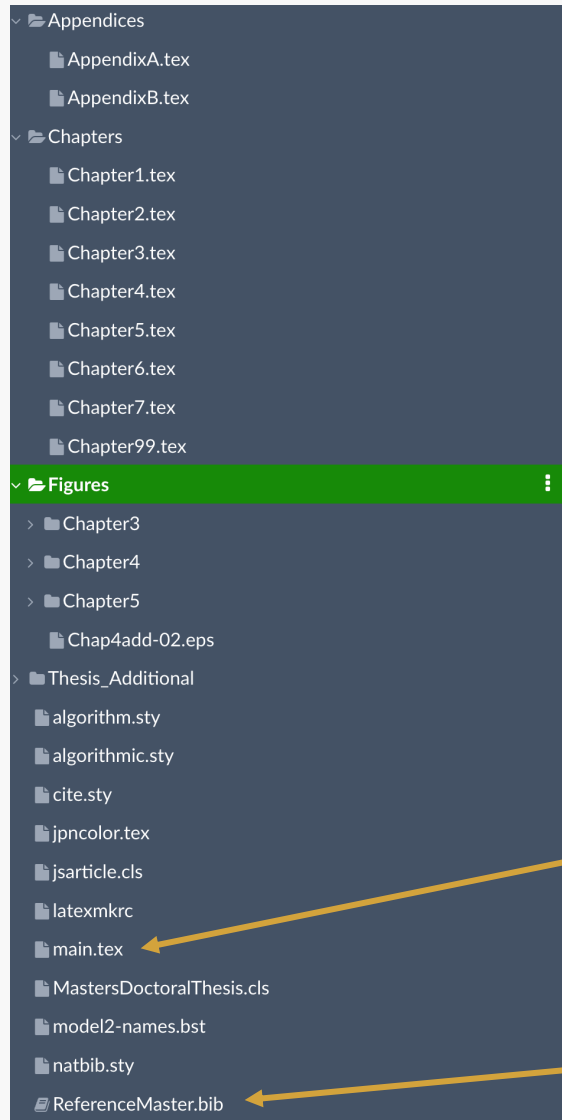
FIGURE 4.8: transfer rho-square for each dataset $v = [1890, 1959]$

To test temporal transferability, we validate estimated models. For the classification and definition of validation in the Behavior model, please refer to Parady et al. (2021). This subsection examines the external validation. However, the samples which compose the dataset are almost the same before and after a particular year. Therefore, the internal validation will be partly examined.

We employ transfer rho-square as performance measures. The transfer rho-square is absolute and easy to use as a benchmark. The definition of transfer rho-square is

開いてる
文書の構成

Tex のフォルダ構成



基本的に1文書1ディレクトリで構成される
オンラインでもオフラインでも同じ

各章のTeX ファイル
本文を書き込むファイル

挿入する画像
章ごとにフォルダでまとめておくと良い

メインのTeX ファイル
Chapetrsディレクトリのファイルを読み込む

bib ファイル 参考文献の情報を保管

Texの入力方法

書いてみよう

- 適当に本文を書いてみる
- 節や項には分かりやすい名前でもラベルをつけておくと、後から呼び出せる
- 呼び出すときは `¥ref{ラベルの名前}` と入力する
(¥はバックslash, Macはoption を押しながら¥)

```
\chapter{Latexを試す}\label{2}
本章の目的はLatex の入力方法に慣れることにある。
なお、昨日の晩御飯は冷凍チャーハンであった。
冷凍チャーハンを含め、冷凍食品の品質は年々向上していると言える。

\section{数式の書き方}
```

Texの入力方法

数式の書き方

```
1 \chapter{Latexを試す}\label{2}
2 \section{数式の書き方}
3
4 本文中に書く場合 $a+b=c$ と書く
5
6 \begin{equation}
7 x^2+y^2 = r^2 \label{circle}
8 \end{equation}
9
10 ラベルを振っておくと
11 数式\ref{circle}と引用することができる
12
13 \begin{align}
14 y&=ax^2+bx+c \\
15 &=a(x+p)^2+q
16 \end{align}
```

- 本文に書く場合
\$記号で囲う
- 改行して書く場合
equation 環境を使う, ラベルを付けること
で参照を可能にする
- 複数行の数式を書く場合
align 環境を使う, $\&$ で改行, $\&$ で場所
を揃える

Texの入力方法

数式の書き方：ビルド

```
1 \chapter{Latexを試す}\label{2}
2 \section{数式の書き方}
3
4 本文中に書く場合 $a+b=c$ と書く
5
6 \begin{equation}
7 x^2+y^2 = r^2 \label{circle}
8 \end{equation}
9
10 ラベルを振っておくと
11 数式\ref{circle}と引用することができる
12
13 \begin{align}
14 y&=ax^2+bx+c \\
15 &=a(x+p)^2+q
16 \end{align}
```

2.1 数式の書き方

本文中に書く場合 $a + b = c$ と書く

$$x^2 + y^2 = r^2 \quad (2.1)$$

ラベルを振っておくと数式 2.1 と引用することができる

$$y = ax^2 + bx + c \quad (2.2)$$

$$= a(x + p)^2 + q \quad (2.3)$$

Texの入力方法

数式の書き方

```
\begin{equation}
G(y)=\sum_m\alpha_m(\sum_k y_k^{1/\mu_m})^{\mu_m}
\end{equation}
```

上付き添字は^
二文字以上なので{}で囲っている

シグマ ギリシャ文字 下付き添字は_

$$G(y) = \sum_m \alpha_m \left(\sum_k y_k^{1/\mu_m} \right)^{\mu_m} \quad (2.4)$$

練習問題

問 1 : MNL モデル

$$P(i) = \frac{\exp(\mu V_i)}{\sum_{i'} \exp(\mu V_{i'})}$$

問 2 : GEV モデル

$$G(\mathbf{y}) = \sum_{m=1}^M a_m \left(\sum_{i \in \mathcal{B}} y_i^{\frac{1}{1-\sigma_m}} \right)^{1-\sigma_m}$$

問 3 : RL モデル

$$V^d(k) = \begin{cases} \mu \log \sum_{a \in A(k)} \exp \left(\frac{1}{\mu} (v(a | k) + \beta V^d(a)) \right) & (k \neq d) \\ 0 & (k = d) \end{cases}$$

Texの入力方法

図の挿入

```
\begin{figure}[h]
  \begin{center}
    \includegraphics[width=141mm]{Figures/Chapter3/Chapter4-05.eps}
    \caption{Comparison allocation parameter concept of CNL, SCL, GNL, GSCL model}
    \label{fig:chap4-1}
  \end{center}
\end{figure}
```

位置の指定

図のセンタリング

図のサイズ(横幅141mm)

図のパス

ラベル

図のキャプション

図の位置はLatexが適当に選ぶか、1行目の[]に優先したい順番に並べる

例：¥\begin{figure}[htb]

h : here (その場所)

b : bottom (ページ下部)

t : top (ページ上部)

p : page (独立したページ)

Texの入力方法

表の挿入

```
\begin{table}[t] ← 位置の指定
  \begin{center}
    \caption{Backward induction algorithm \citep{oyama2017discounted}} ← 表のキャプション
    \label{tab:chap3backward} ← ラベル
    \begin{tabular}{lll} ← 列ごとに文字列の位置指定
      \toprule
      \textit{Step1} & set &  $d=s_T, V^d_{t,c}(s_T)$  \\
      \textit{Step2} & initialization &  $t=T, V^d_{c}(s_{\tau})=0$  \\
      \textit{Step3} & backward calculation &  $t=\tau-1$  and calculate  $V^d_{c}(s_{\tau})$  \\
      \textit{Step4} & final decision & if  $t=0$ , end. Otherwise, execute Step 3. \\
      \bottomrule ← 横罫線
    \end{tabular}
  \end{center}
\end{table}
```

改行

tabular環境設定時に、列数の指定と列ごとの文字の位置を指定する

l : 左寄せ c : 中央揃え r : 右寄せ | : 縦罫線

Texの入力方法

参考文献の挿入

- bib ファイルを使う

```
@article{mai2018decomposition,  
  title={A decomposition method for estimating recursive logit based route choice  
  models},  
  author={Mai, Tien and Bastin, Fabian and Frejinger, Emma},  
  journal={EURO Journal on Transportation and Logistics},  
  volume={7},  
  number={3},  
  pages={253--275},  
  year={2018},  
  publisher={Elsevier}  
}
```

- 本文では`\cite` を使って呼び出す

- bib ファイルの書き方

- Google Scholar を使って引用したい論文の「引用>BibTeX」でBibTeX形式の引用をコピーする

良い作図をするために

1. 3D は極力やめる

まずは2D で表現できないか考えましょう

2. 要素を極力少なくする

情報量ゼロの要素はないか確認しましょう

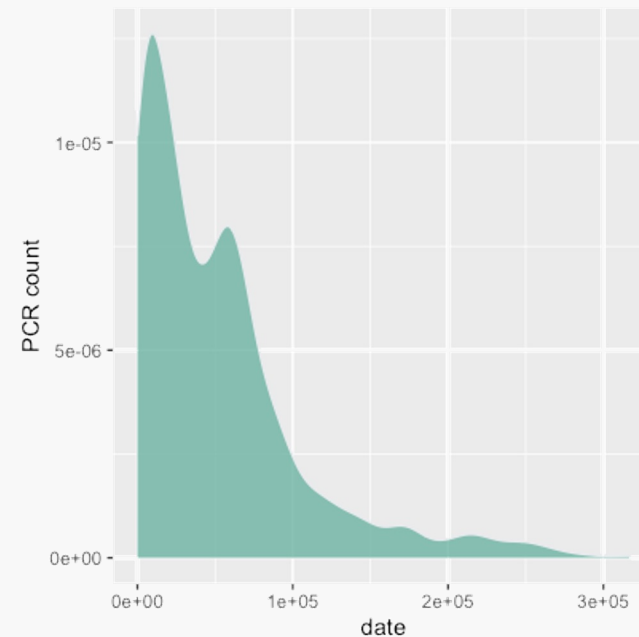
右図の場合, 薄いグレーの背景である必然性はありません

緑色である必然性もありません

グラフの目盛り線も本当に必要か検討しましょう

3. モノクロ印刷を前提にする

色盲や色弱の人への配慮を忘れずに ユニバーサルな色使いをしましょう



R言語とは

- 統計処理に特化したプログラミング言語と環境
- オープンソース・フリーソフトウェア
- 豊富な「パッケージ」を利用して、容易に機能拡張が可能
- Tipsがネットに豊富
- 特徴
 - データの分類・集計・整理が得意
 - ベクトル、行列演算が簡単
 - 初学者に優しい
 - グラフ等を使った可視化ができる

R言語とは

- RとR studio をインストールする
- R本体 (Mac)
 - <https://cran.r-project.org/bin/macosx/>
 - Intel mac とApple silicon mac とではインストールするパッケージが違うので注意
- Rstudio
 - <https://www.rstudio.com/products/rstudio/download/>
 - Rstudio 社が開発したRの総合環境開発
 - Free のRstudio Desktop でOK

Rstudio の構成

The screenshot displays the RStudio environment with three main components highlighted by yellow dashed boxes and labels:

- エディタ (Editor):** The top-left pane shows an R script with code for loading packages, setting directories, and running a simulation loop. The code includes comments in Japanese and R syntax for file paths and data manipulation.
- コンソール (Console):** The top-right pane shows the execution output, including numerical results for variables like 'b', 'l', 'dl', and 'B', and a warning message about file monitoring.
- 作業環境内の変数 (Variables in the workspace):** The bottom-right pane shows a list of variables and their values, such as 'b', 'dire', 'input', 'input_area', and 'iter'.

At the bottom, the **ファイルグラフィックスパッケージヘルプ (File Graphics Package Help)** pane shows a file browser view of the current project directory, listing various R script files and their sizes and modification dates.

Rstudio の使い方

```
> 1+2*3/4 # 四則演算ができます. #をつけるコメントアウトできます  
[1] 2.5
```

```
> 3^2 # 累乗は^  
[1] 9
```

```
> x<-6 # 変数の代入は<-  
> print(x) # コンソール上に表示  
[1] 6
```

```
> y<-7  
> x*y # 変数同士の計算  
[1] 42
```

```
> rm(list=ls()) # 変数を全て削除
```

Rstudio の使い方 : ベクトル

```
> v1 <- c(1,2,3,4,5) # 数字を鎖状につなげ行ベクトルを作る
> v1
[1] 1 2 3 4 5
> v2 <- 6:10 # 左の番号を最小値, 右の番号を最大値として取る
> v2
[1] 6 7 8 9 10
> v3<-15:11 # 逆順でもOK
> v3
[1] 15 14 13 12 11
> v4<-seq(11,15, by=1) # 一つ目の引数から二つ目の引数まで1ずつ増える
> v4
[1] 11 12 13 14 15
> v5 <- rep(2,5) # 2を5つ繰り返す
> v5
[1] 2 2 2 2 2
```

Rstudio の使い方 : ベクトル

```
> v1[1] # ベクトルの1番目の要素にアクセス. pythonやCと違い, 1から数え始める
[1] 1

> v1[2:3] # 左の数字から右の数字まで番目の要素にアクセス
[1] 2 3

> v1[c(2,4)] # 複数の要素にもアクセスできる
[1] 2 4

> v1[-1] # マイナスはそれ以外
[1] 2 3 4 5

> length(v1) # ベクトルの長さを計算
[1] 5
```

Rstudio の使い方 : ベクトル

```
> v1+v3      # ベクトル同士の演算
[1] 16 16 16 16 16

> v1*v3      # ベクトル要素同士の掛け算
[1] 15 28 39 48 55

> v1%*%v3    # ベクトルの内積
      [,1]
[1,] 185

> v1+2       # ベクトルの各要素に2が足される
[1] 3 4 5 6 7

> X1 <- c(v1,v2) # 行ベクトル同士が結合
> X1
> [1] 1 2 3 4 5 6 7 8 9 10
```

Rstudio の使い方 : 行列

```
> L1 <- matrix(X1, nrow=2, ncol=5) # nrowで行数, ncolで列数を指定
> L1
      [,1] [,2] [,3] [,4] [,5]
[1,]    1    3    5    7    9
[2,]    2    4    6    8   10
> L2 <- rbind(v3,v4) # 行ベクトル単位で結合, 列ベクトル単位の場合はcbind
> L2
      [,1] [,2] [,3] [,4] [,5]
v3     15   14   13   12   11
v4     11   12   13   14   15

> ncol(L1) # 列数
[1] 5

> nrow(L1) # 行数
[1] 2
```

Rstudio の使い方 : 行列

```
> L1[2,3]      # 2行目3列目の要素にアクセス
[1] 6

> L1[2,]       # 2行目の要素にアクセス
> [1] 2  4  6  8 10

> L1[,1]       # 1列目の要素にアクセス
> [1] 1 2 3

> rowSums(L1)   # 行の総和
[1] 25 30

> colSums(L1)   # 列の総和
[1]  3  7 11 15 19
```

Rstudio の使い方 : 行列の計算

```
> L1+L2 # 行列の要素同士の足し算
```

```
  [,1] [,2] [,3] [,4] [,5]  
v3   16  17  18  19  20  
v4   13  16  19  22  25
```

```
> L1*L2 # 行列の要素同士の掛け算
```

```
  [,1] [,2] [,3] [,4] [,5]  
v3   15  42  65  84  99  
v4   22  48  78 112 150
```

```
> L1 %*% t(L2) # t()は転置行列, %*%は内積
```

```
  v3  v4  
[1,] 305 345  
[2,] 370 410
```

Rstudio の使い方 : 行列の計算

```
> diag(3)      # 3*3の単位行列を作る
      [,1] [,2] [,3]
[1,]    1    0    0
[2,]    0    1    0
[3,]    0    0    1

> L3<- matrix(c(2,1,1,2),2,2)
> solve(L3)     # 逆行列を求める
      [,1]      [,2]
[1,]  0.6666667 -0.3333333
[2,] -0.3333333  0.6666667
```


Rstudio の使い方 : 行列の計算

```
> apply (L1, 1, sum) # 各行の和, colSums(x)と同じ  
[1] 25 30
```

```
> apply (L1, 1, mean) # 各行の平均  
[1] 5 6
```

```
> apply (L1, 2, sum) # 各列の和, rowSums(x)と同じ  
[1] 3 7 11 15 19
```

```
> apply (L1, 2, mean) # 各列の平均  
[1] 1.5 3.5 5.5 7.5 9.5
```

```
# apply 関数の3つ目の引数には自作関数を入れても良い
```

Rstudio の使い方 : 関数

```
> Square <- function(x) { # x は代入する変数. 2変数でも良い
+   X2 <- x*x
+   return(X2)           # return で返す値を定義
+ }
> Square(2)
[1] 4

> Square(v1)             # ベクトルを代入してもOK
[1] 1 4 9 16 25

> Square(L1)            # 行列を代入してもOK
      [,1] [,2] [,3] [,4] [,5]
[1,]    1    9   25   49   81
[2,]    4   16   36   64  100

> apply(L1, 1, Square)  # apply関数は自作関数を適用することができる
```

Rstudio の使い方 : 演算子

```
> 1==2          # == 等しい
[1] FALSE
> 1!=2          # != 等しくない
[1] TRUE
> 1<=2         # >= は ≥, <= は ≤
[1] TRUE
> 1>2
[1] FALSE
> x<-7
> (x > 2) && (x < 4) # 「xは2より大きい」かつ「xは4より小さい」
[1] FALSE
> (x > 2) || (x < 8) # 「xは2より大きい」または「xは8より小さい」
[1] TRUE
```

Rstudio の使い方 : if文

```
> x<-4
> if(x<0){
+   print("0より小さい")
+ } else{
+   print("0以上")
+ }
[1] "0より大きい"
```

if(条件式){式}で条件分岐

```
> if(x<0){
+   print("0より小さい")
+ } else if(x>5){
+   print("5より大きい")
+ } else{
+   print("0以上5以下")
+ }
[1] "0以上5以下"
```

三つ以上の条件分岐には, else if を使う

Rstudio の使い方 : for文とwhile文

```
> x <- 0
> for (i in 1:10) {      # 繰り返し文for(ループ変数 in ベクトルやリスト)
+   x <- x+i
+ }
> print(x)
[1] 55
```

```
> x <- 0
> while (i <= 10) {     # while(条件式)
+   x <- x+i           # 条件式がTRUE である限り式が繰り返される
+   i <- i+1
+ }
> print(x)
[1] 10
```

Rstudio の使い方 : for文とwhile文

```
> x <- 0
> for(i in 1:10){
+   x <- x+1
+   if (x == 3) break # x==3 がTRUEになったら
+ }                 # for文から抜ける
> x
[1] 3
```

```
> x <- 0
> repeat {
+   if (x <= 5) x <- x+1 # break 文が見つかるまで式が繰り返される
+   else break # x > 5 ならば repeat 文から抜ける
+ }
> x
[1] 6
```

Rstudio の使い方 : グラフの作成

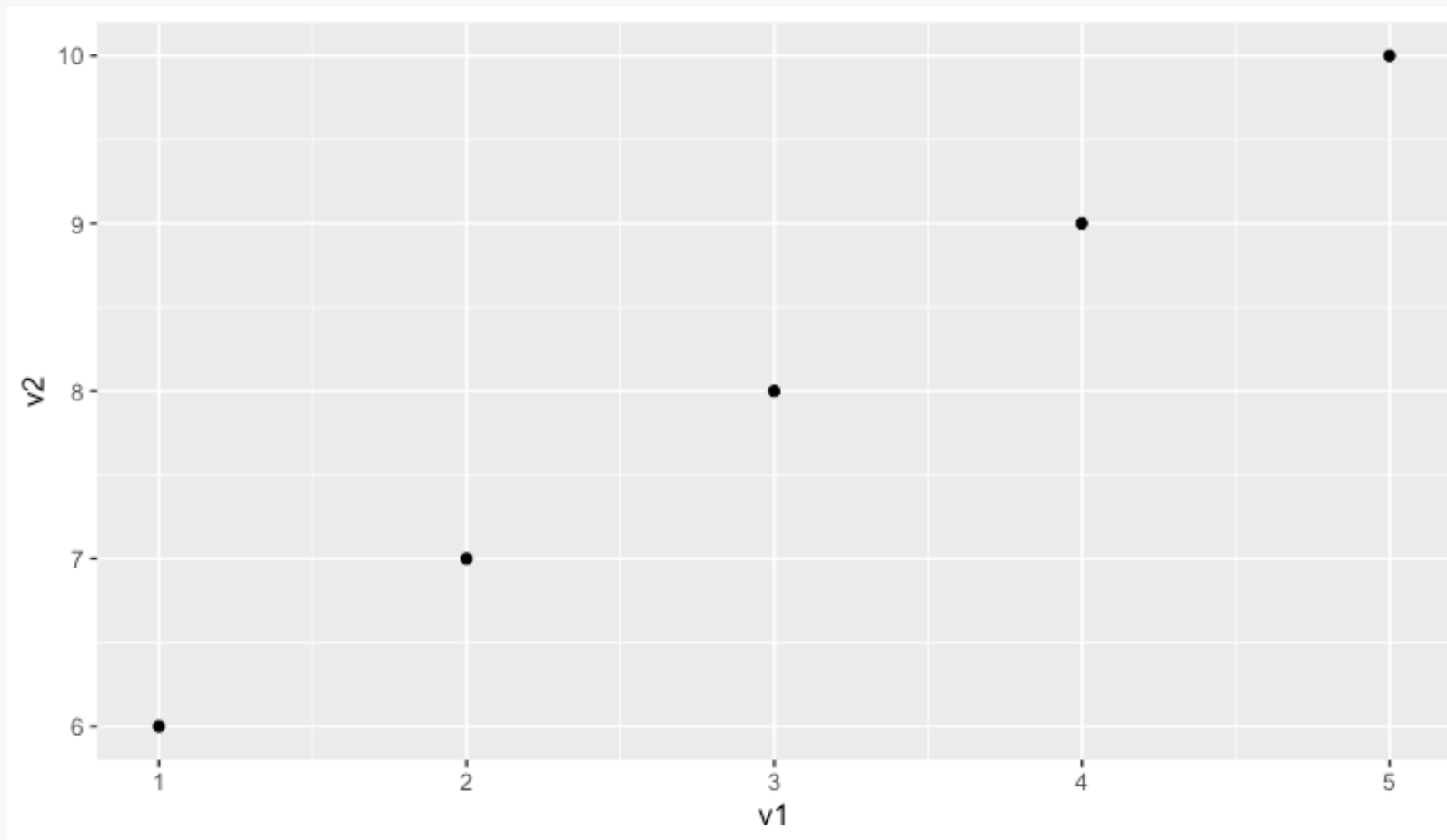
```
> install.packages("ggplot2")
> install.packages("dplyr")
> library(ggplot2) # 描画用のパッケージを読み込む
> library(dplyr)   # データ操作をするパッケージを読み込む

> v1
[1] 1 2 3 4 5

> v2
[1] 6 7 8 9 10
> data<-data.frame(x=v1,y=v2) # データフレームを生成
> data %>%                    # グラフの元になるデータをパイプ演算子 %>% で繋ぐ
+   ggplot() +
+   geom_point(aes(x=x, y=y))+ # 幾何オブジェクトpoint を選択
+                               # aesでx軸の要素とy要素を定義する
+   labs(x = "v1", y="v2")
>
```

Rstudio の使い方 : グラフの作成

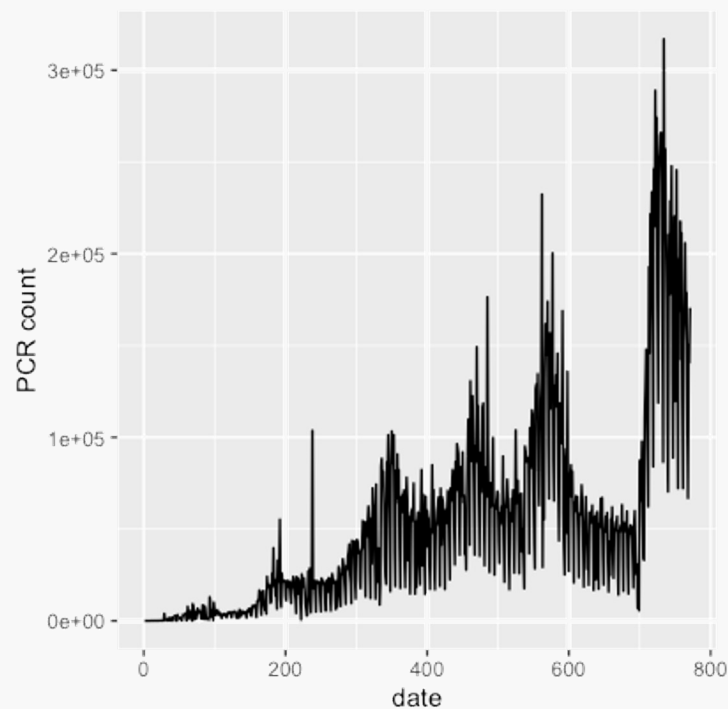
- 描画の結果, 以下のグラフが出力されます



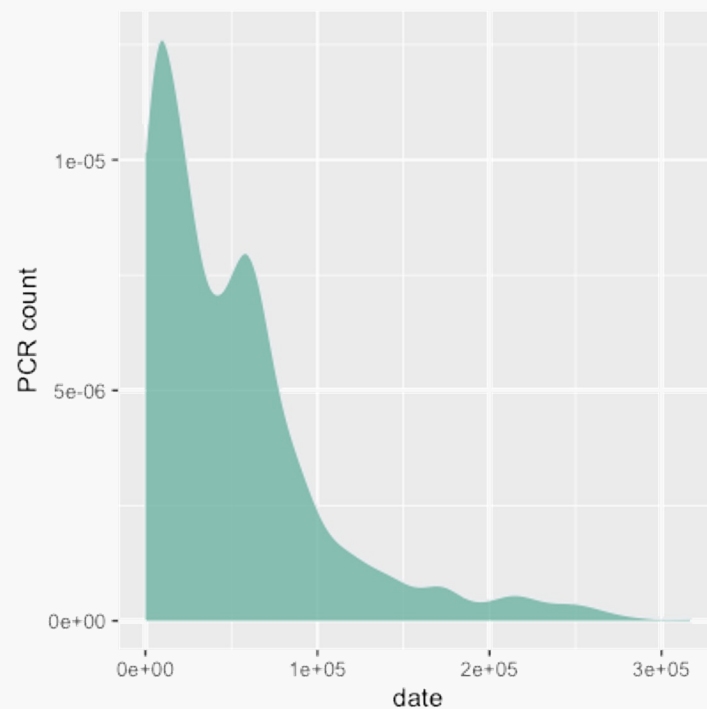
Rstudio の使い方 : グラフの作成

- 幾何オブジェクトの種類

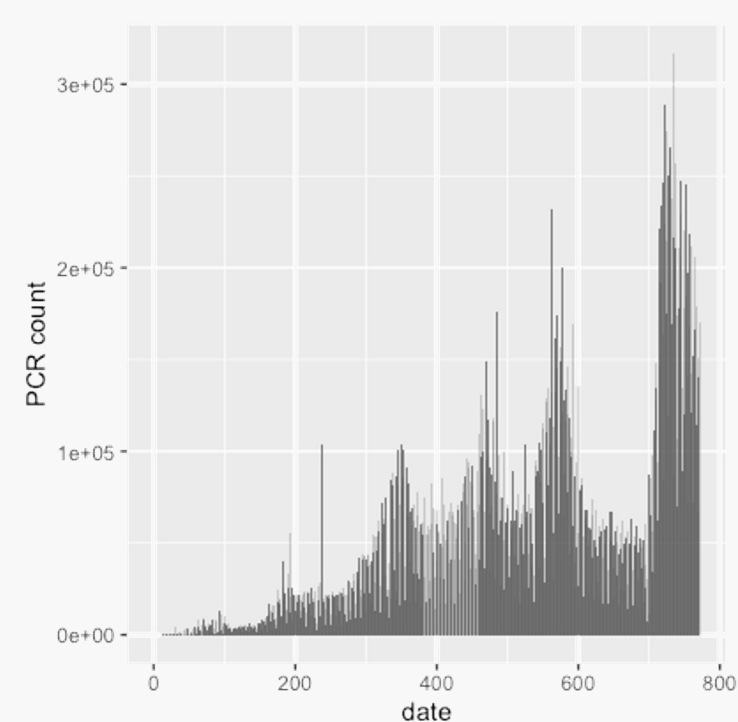
geom_line



geom_density



geom_bar



データはPCR 検査数 : <https://www.mhlw.go.jp/stf/covid-19/open-data.html>

Rstudio の使い方 : データ処理

data.frame

- 数値ベクトル、文字ベクトル、因子ベクトルなど異なる型のデータをまとめて1つの変数として持つdata.frameという型が統計処理に便利。
- 各行・列にラベルをつけることができ、ラベルによる操作が可能。

作業ディレクトリ

- ファイルからデータやプログラムを読み込んだり、ファイルに書き込んだりする場所

```
# データファイル"pcr_tested_daily.csv"を読み込み
data <- read.csv("pcr_tested_daily.csv", stringsAsFactors=F, encoding="UTF-8")

# 読み込んだdata.frame の列名を"date"と"count"に変更
names(data) <- c("date", "count")
```

データはPCR 検査数 : <https://www.mhlw.go.jp/stf/covid-19/open-data.html>

Rstudio の使い方 : データ処理

```
data[100,] #100行目を表示
```

```
  date      count  
100 2020/5/14  5866
```

```
> data$date # 列名を指定して「date」列を表示
```

```
[1] "2020/2/5"    "2020/2/6"    "2020/2/7"    "2020/2/8"    "2020/2/9"  
[6] "2020/2/10"   "2020/2/11"   "2020/2/12"   "2020/2/13"   "2020/2/14"  
[11] "2020/2/15"  "2020/2/16"  "2020/2/17"  "2020/2/18"  "2020/2/19"  ...
```

```
data[data$date=="2021/1/15",] # 2021/1/15の検査数を表示
```

```
  date      count  
346 2021/1/15 70610
```

Rstudio の使い方 : データ処理

```
# 検査数が10万以上の日(date)を抜き出す
> data$date[data$count>100000]
 [1] NA          NA          NA          "2020/9/29"
 [5] "2021/1/14" "2021/1/19" "2021/1/22" "2021/5/7" ...

# 2022/01/01以降で検査数が20万件以上の日と検査数を抜き出す
> data[(data$date>="2022/01/01")&(data$count>200000),]
      date  count
715 2022/1/19 221682
717 2022/1/21 233676 ...

install.packages("lubridate")
library(lubridate) # 日付を丸めるパッケージ
data$month <- floor_date(as.Date(data$date), "month") # 月ごとに丸めて「month」列に代入
```

Rstudio の使い方 : データ処理

```
# 月ごとの検査数の1日あたり平均を知りたい
```

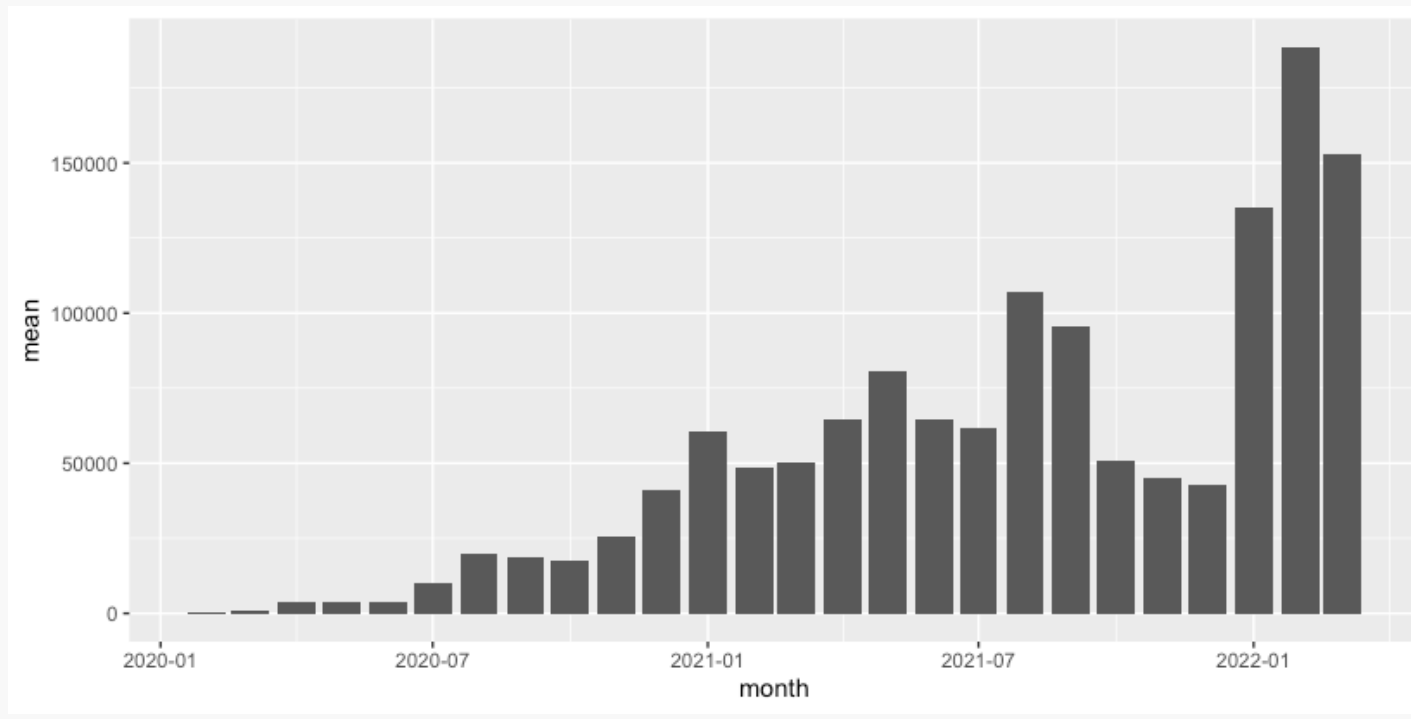
```
> data %>%  
+   group_by(month) %>% # 月ごとにグループ化  
+   summarize(mean=mean(count, na.rm=T)) # 月ごとの平均を計算  
# A tibble: 26 x 2  
  month          mean  
  <date>        <dbl>  
1 2020-02-01     62.4  
2 2020-03-01  1045.  
3 2020-04-01  3943. ...
```

```
# 月ごとの検査数の1日あたり平均を図化する
```

```
data %>%  
  group_by(month) %>% # 月ごとにグループ化  
  summarize(mean=mean(count, na.rm=T)) %>%  
  ggplot(aes(x=month, y=mean)) +  
  geom_bar(stat = "identity")
```

Rstudio の使い方 : データ処理と描画

```
# 月ごとの検査数の1日あたり平均を図化する
data %>%
  group_by(month) %>% # 月ごとにグループ化
  summarize(mean=mean(count, na.rm=T)) %>% # summarize で集計, グループごとの平均
  ggplot(aes(x=month, y=mean)) + # x軸にmonth, y軸に平均をとる
  geom_bar(stat = "identity")
```



研究発表のルール ～ いい発表のために

1. 聞き手のことを考えましょう

話には関係のない図がないか、逆に足りない図がないかチェックしましょう

2. エクスキューズはやめましょう

するくらいなら盛り込みましょう、厳しいですが言い訳は聴く気を削ぎます

上手く盛り込めなかったり悩んでいることがある時は最後にまとめて話すのが互いに吉

3. モノクロ印刷を前提にしましょう

色盲や色弱の人への配慮を忘れずに ユニバーサルな色使いをしましょう

4. 録画してチェックしましょう

耐え難いと思いますが、改善効果はてきめんです

まとめ

- 色んなツールを使って研究を進めることになります
- RやTex, Python でわからないことがあれば検索すること
 - ネットにたくさん落ちてます
 - 先人の蓄積はどんどん有効活用しましょう
- わからないことは先生，先輩，同期に遠慮なく聞きましょう
 - SOS は早ければ早いほど良いです
- 聞かれたら一緒に考えましょう
 - 聞かれたことが仮に分からなくても，自分の糧になる
 - 逆に分からないと言われても責めない
 - 遠慮なく聞きあえる雰囲気と関係は良い(研究)コミュニティづくりの基本です