バイオスタティスティクス基礎論 ガイダンス

岩田洋佳 hiroiwata@g.ecc.u-tokyo.ac.jp

2021/4/9

講義の内容

バイオスタティスティクス基礎論は、R と RStudio を用いて、実習形式で統計の初歩を 学びます。R は統計解析のためのコンピュータ言語で、RStudio はそれを利用するための ソフトウエアの一つです。R をインストールするだけでも、R 言語を使うことができます が、本講義では RStudio を用いて授業を行います。

講義で扱う内容は、以下のとおりです。

- 1. R でデータを視覚化する(4月9日)
- 2. 回帰分析、分散分析(4月16日)
- 3. 主成分分析、多次元尺度構成法(4月23日)
- 4. 階層的クラスタ解析、非階層的クラスタ解析(4月30日)

授業の進め方

ファイルは事前にダウンロードしてください

各回の授業は、いずれも、テキスト(PDF)、Rコード(R)、データ(CSV)ファイル を用いて実習形式で進めていきます。全てのファイルについて、遅くとも講義の前日まで には、アグリバイオのホームページおよび ITC-LMS からダウンロードできるようにしてお きます。講義の前に必ずダウンロードを行っておいて下さい。

授業の途中でクイズを行う予定です

授業に、少しメリハリやゲーム性をつけるために、途中でクイズを実施します。ここで も、少し練習をしてみましょう。

- 1. まず、https://www.menti.com/95twm188to に移動して下さい。
- 2. 次に、ニックネームを入れて、"Join quiz"をクリックして下さい。

	≅ menti.com Č		<u> </u>
Enter your name			
Hiroyoshi		21	
	Join quiz		
Powered	by Mentimeter Terms		

3. すると待機状態になります。しばらくするとクイズがはじまります。皆さんの健闘を 祈ります。



なお、このクイズは一種のゲームで、成績評価には一切関係しません。楽しんで取り組ん で下さい。

成績の評価

成績の評価は、出席(ITC-LMSを用いて出席をとる予定です)と毎回出されるレポート(ITC-LMSを用いて提出してもらう予定です)をもとに評価します。

R と RStudio のインストール

Windows の場合

Rのインストール

- 1. まず、CRAN のサイト https://cran.ism.ac.jp/ に移動します。
- 2. "Download R for Windows"をクリックします。



3. "base"を選択します。



4. "Download R 4.0.5 for Windows"をクリックします。最新バージョンはそのときによっ て異なります。



5. ダウンロードをするときは、「実行」ではなく、「保存」をして下さい。

6. ダウンロードして保存されたファイルをクリックして、R をインストールします。な お、必ず「管理者として実行」(管理者権限でインストール)してください。



7. インストール時のオプションは、特に変更をする必要はありません。デフォルトの設定のままでインストールを進めていきます。

R Studio のインストール

- 1. RStudio のダウンロードサイト https://rstudio.com/products/rstudio/ に移動します。
- 2. 次に、"RStudio Desktop"をクリックします。



3. "DOWNLOAD RSTUDIO DESKTOP"をクリックします。

$\leftarrow \rightarrow 0$	<u>ن</u> ۵	https://rstudio.com/products/rstudio/	0 🛧 🖕	¥ 1	L B	
		 diagnose and fix errors quickly Extensive package development tools 				-
	Support	Community forums only	 Priority Email Support 8 hour response during business hours (ET) 			ł
	License	AGPL v3	RStudio License Agreement			
	Pricing	Free	\$995/year			
		DOWNLOAD RSTUDIO DESKTOP	DOWNLOAD FREE RSTUDIO DESKTOP PRO TRIAL			
			Purchase Contact Sales			
https://rstudio.com/	/products/rstudic	/download/				

4. 一番左(フリー版の RStudio Desktop)の"DOWNLOAD"をクリックします。



- 5. ダウンロードをするときは、「実行」ではなく、「保存」をして下さい。
- **6.** ダウンロードして保存されたファイルをクリックして、**RStudio** をインストールしま す。なお、必ず「管理者として実行」(管理者権限でインストール)してください。



インストール時のオプションは、特に変更をする必要はありません。デフォルトの設定のままでインストールを進めていきます。

Mac OS の場合

Rのインストール

- 1. まず、CRAN のサイト https://cran.ism.ac.jp/ に移動します。
- 2. "Download R for (MAC) OS X"をクリックします。



3. 次の図の左下にある"R-4.0.5.pkg"をクリックします。最新バージョンはそのときによって、異なります。



4. ダウンロードして保存されたファイルをクリックして、R をインストールします。



5. インストール時のオプションは、特に変更をする必要はありません。デフォルトの設定のままでインストールを進めていきます。

R Studio のインストール

- 1. RStudio のダウンロードサイト https://rstudio.com/products/rstudio/ に移動します。
- 2. 次に、"RStudio Desktop"をクリックします。



3. "DOWNLOAD RSTUDIO DESKTOP"をクリックします。

•• <>		📰 🔒 rstudio.com	Ċ	00
	R Stu	dio Desktop		
		Open Source Edition	RStudio Desktop Pro	
	Overview	Access RStudio locally Syntax highlighting, code completion, and smart indentation Execute R code directly from the source editor Quickly jump to function definitions Easily manage multiple working directories using projects Integrated R help and documentation Interactive debugger to diagnose and fix errors quickly Extensive package development tools.	All of the features of open source; plus: A commercial license for organizations not able to use AGPL software Access to priority support RSoudio Professional Drivers	
	Support	Community forums only	Priority Email Support Bhour response during business hours (ET)	
	License	AGPL v3	RStudio License Agreement	
	Pricing	Free	\$995/year	
		DOWINE, OAD INSTUDIO DESISTOP	DOWNLOAD FREE RETUDIO DESKTOP PRO TRIAL	

4. 一番左(フリー版の RStudio Desktop)の"DOWNLOAD"をクリックします。

		i rstudio.com	Ċ	ð Ø
	RStudio Desktop Open Source License	RStudio Desktop Commercial License	RStudio Server	RStudio Server Pro
	Free	\$995 /year	Free	\$4,975 /year (S Named Users)
	DOWNLOAD	BUY	DOWNLOAD	BUY
	Learn more	Learn more	Learn more	Evaluation Learn more
Integrated Tools for R	×	×	~	~
Priority Support		~		~
Access via Web Browser			~	~
Enterprise Security				~
Project Sharing				×
Manage Multiple R Sessions & Versions				~
Admin Dashboard				×
Load Balancing				~
Auditing and Monitoring				~

5. ダウンロードして保存されたファイルをクリックすると以下のように表示されます。 あとは、RStudio をドラックして、Applications の上にドロップします。

• • •	RStudio-1.2.5033	
,A.	R	
Applications	RStudio	
- appressions	no baare	

RとRStudio を使ってみる

RとRStudio が動くことを確認する

ここからは、Windows と Mac OS を一緒にして説明していきます。それは、RStudio のイ ンターフェースが Windows と Mac OS で一緒だからです。

まずは、左側の Console に、次のように打ち込んでみましょう。

3 + 4 * 2

[1] 11



計算が実行され、結果が返ってきます。R の使い方については、来週からの授業で詳しく 説明するのでこのくらいにしておきましょう。

RとRStudioを講義で使うための準備をする

R のパッケージをインストールする

R は、そのままでもいろいろな統計解析を実施することができますが、パッケージとよば れる「拡張セット」のようなものをインストールすると、利用できる解析手法が大幅に広 がります。ここでは、パッケージのインストールの仕方について説明します。なお、イン ストールを行うためには、パソコンがインターネットに接続されていなければなりません。

1. RStudio のメニューの Tools から、"Install Packages"を選びます。

Ś	RStudio	File	Edit	Code	View	Plots	Session	Build	Debug	Profile	Tools	Window	Help	
											Insta	II Packages	i	
											Chec	k for Packa	ige Update	s
											Versi	on Control		•
											Shell Term Jobs Addir	 inal 15		* * *
											Keyb Modi	oard Short fy Keyboar	cuts Help d Shortcuts	ток
											Proje	ct Options		公36,
											Globa	al Options.		ж,

2. そのあと示されるダイアログボックス(下図中央)の Packages に、インストールしたいパッケージ名を入力します。ここでは、plotly というパッケージをインストールしてみます。



3. plotly パッケージがダウロードされてインストールされます。



4. 以下のコマンドを Console に打ち込んで、図を描いてみます。

library(plotly)

plot_ly(midwest, x = ~percollege, color = ~state, type = "box")

RStudio で実行すると、このように表示されます。



Rの Working Directory を設定する

R では、R を実行する「作業場」"Working Directory"を指定しなければならない場合が 少なくありません。これは、授業で使う R コードやデータを R が見つけられるようにする ためです。ここでは、RStudio を使って、これを指定してみましょう。

まず最初に、以下の Google Drive から、使用する R コード(readData.R) とデータ (iris.csv) をダウンロードします。

https://drive.google.com/drive/u/2/folders/160TTfRf0Lz9qxm8vLdciiaYZmS0Zr9BL

次に、これらのファイルをどこか自分の作業しやすいフォルダ(ディレクトリに移動して みましょう)。ここでは、デスクトップに Rtemp というフォルダを作って保存したとしま す。

RStudio で、このフォルダを"Working Directory"として指定するには、以下のようにします。

1. RStudioの右側下にある「Files」というタブをクリックします。

🕒 • 👒 🖙 • 📄 💮 📥 🗚 Coto Ne/function 👘 🔛 • Addins +			Project: (None)
madData.R ×	Environment History Connections		-0
🗇 🗇 📊 🖂 Source on Save I 🔍 🎢 + 🔄 🗃 Run I 🍽 🕒 Source + 🗟	😅 📊 🖙 Import Dataset + 🖌		💷 Ust + 🛞 -
1 dota <- read.csv("Lris.csv", row.names = 1) 2 head(data) 3	n Global Environment +		Q,
4 # preparation for the next week 5 required.packages <- c("KernSmooth", "maps", "mapdata", "plot! 6 new.packages <- required.packages[[(required.packages %in% ins] 7 x ((Cleanth/new.packages)). / 8 ii (Topteret) = RSonge = RSonge =	Environment is e	mpéy	
Console Terminal × Jobs ×			
-10			
<pre>Copyright (C) 2820 The R Foundation for Statistical Computing Flatform: x86,64-apple-dorwin15.6.0 (64-bit) R は、自由なソフトウェアであり、「完全性解理」です。 一定の条件に気は、自由にこれを再整合することができます。 整布条件の詳細に関しては、'license()' あるいは 'licence()' と入力してくださ い。 R は多くの員動者による共同プロジェクトです。 詳しくは 'contributors()' と入力してください。 また、R や R のパッケージを出版物で引用する際の形式については 'citotion()' と入力してください。 'demo()' と入力すれば不至をあることができます。 'help.start()' で HTML ブラウザによるヘルプがみられます。 'g()' と入力すれば R を相了します。 ></pre>	Hiss OS Packages Prop Viewer Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image: Annual Stress Image	5278 8 282 8 780.5 KB	Apr 9, 2020, 10:38 PM Apr 9, 2020, 12:34 AM Apr 9, 2020, 10:38 PM

- 2. 「Home > Desktop」となっているところや、その下に示されているフォルダをクリッ クすることで、パソコン内にある様々なフォルダに移動できます。
- 3. ここでは、「Rtemp」に移動したいので、Rtemp をクリックします。保存しておいた R コードやデータのファイルがあることが分かります。
- **4. R** コードやデータがある場所まで移動できたら、歯車型のアイコンの横に「More」と書いてあるところをクリックします。



5. するとドロップダウンメニューが示されますので、「Set As Working Directory」を選びます。

- 👒 🦙 - 🕞 📄 📄 🍌 Co to file function		
		Project: (None)
Console Terminal × R Markdown × Jobs ×	Environment History Connections	- 6
~/Desktop/ 🖗 🦪	🚰 🔒 📑 Import Dataset 🔹 🍕	🗏 List • 🛛 🞯
<pre>~/D43x00/~</pre> ~/D43x00/~ */D43x00/~ * Version 3.6.3 (2020-02-29) "Holding the Windsock" (opyright (5) 2020 The R Foundation for Statistical Computing Platform: x86_64-apple-darwin15.6.0 (64-bit) R は、自由なソフトウェアであり、「完全に無保証」です。 = = の条件に従えば、自由にこれを再発格することができます。 Eを条件の課題に関しては、"licence()' と入力してください。 R は多くの貢献者による共同プロジェクトです。 R しくは 'contributors()' と入力してください。 # にくは 'contributors()' と入力してください。 * demo()' と入力すればデモをみることができます。 help()' と入力すればデモをみることができます。 help()' と入力すればデモをみることができます。 help()' と入力すればアモをみることができます。 help()' と入力すればアモをみることができます。 help()' と入力すればアモをみることができます。 help()' と入力すればアモをみることができます。 exted(~~Desktop~) * etmd(~~Desktop~) * etmd(~Desktop~) * etmd(~~Desktop	Clobal Environment - Civinon Files Plots Packages Help Viewer New Folder O Delete Rename Name Name To readData.R	Acre - Copy Copy To Were - Copy Copy To Were - Copy To Set As Working Directory Show Folder in New Window Show Hidden Files

6. これで、"Working Directory"をデスクトップ上にある Rtemp というフォルダにするこ とができました。

こうすることで、Rからデータファイルを簡単に読み込むことができるようになります。

まずは、R コードを開いてみましょう。右下にある「readData.R」というファイルをクリ ックすると、左上に読み込んだコードが示されます。

🔹 🧐 🚰 • 🕞 📄 🧀 🍂 Go to file/function 👘 🐼 • Addins •	_		Project: (None)
readData.R ×	Environment History Connections		
💷 🗌 💭 🔛 🖸 Source on Save 🔍 🎢 🗉 📄 Run 🍽 📑 Source 📲	🖙 🔒 🖙 Import Dataset = 🛛 🍕		🗏 List • 🛛 🙄
1 data <- read.csv("iris.csv", row.names = 1) 2 head(data) 3	🍓 Global Environment -		Q,
	Environmen	t is empty	
3:1 (Top Level) = R Script	: Eller Storr Beckney Male Vieway		_
	New Folder O Delete Rename More		
version 3.6.3 (2020-02-29) "Holding the Windsock"	Home > Deskton > Rtemp		
opyright (C) 2020 The R Foundation for Statistical Computing	A Name	Size	Modified
latform: x86_64-apple-darwin15.6.0 (64-bit)	1		
は、自由なソフトウェアであり、「完全に無保証」です。	Rhistory	852 B	Apr 9, 2020, 10:54 PM
-定の条件に従えば、自由にこれを再配布することができます。	iris.csv	4.1 KB	Apr 9, 2020, 10:14 PM
6布条件の詳細に関しては、'license()' あるいは 'licence()' と入力してください。	eadData.R	55 B	Apr 9, 2020, 10:52 PM
は多くの貢献者による共同プロジェクトです。 #しくは 'contributors()' と入力してください。 #た、R や R のバッケージを出版物で引用する開の形式については citation()' と入力してください。			
demo()、と入力すればデモをみることができます。 help()、とすればオンラインヘルブが出ます。 help.stort()、で HTML ブラウザによるヘルブがみられます。 q()、と入力すれば R を終了します。			

このうち、最初の2行を選択して、実行してみましょう。実行は、Run というアイコンを クリックします。あるいは、Mac OS では「command + return (enter)」、Windows では 「cntl + enter (return)」でも実行できます。

🔹 👒 🚅 • 🔄 🔄 📥 🔿 Go to file/function 🔢 🗟 • Addins •		Project: (None
eadData.R ×	Environment History Connections	-
💷 🗐 🔛 🖸 Source on Save 🛛 🔍 🎢 📲 👄 Run 🛛 🐤 🕒 Source	🔹 🧟 🚰 📊 🐨 Import Dataset 🔹 🍕	🗏 List • 🛛 🙆
1 data <- read.csv("iris.csv", row.ne <u>res = 1)</u> 2 head(data) 3	diobal Environment -	Q.
3:1 (Top Level) 0	Script \$	
Console Terminal × Jobs ×	Files Plots Packages Help View	ver
~/Desktop/Rtemp/ A	New Folder O Delete Rename	🍄 More +
opyright (C) 2020 The R Foundation for Statistical Computing	Home > Desktop > Rtemp	Size Modified
latform: x86_64-apple-darwin15.6.0 (64-bit)	1 marine	2156 10001160
は、自由なソフトウェアであり、「完全に個保証」です。	Bhistory	852 B Apr 9, 2020, 10:54 PM
-定の条件に従えば、自由にこれを再配布することができます。	iris.csv	4.1 K8 Apr 9, 2020, 10:14 PM
?布条件の詳細に関しては、'license()' あるいは 'licence()' と入力してく	Edu, PreadData.R	55 B Apr 9, 2020, 10:52 PM
は多くの貢献者による共同プロジェクトです。		
FUくは 'contributors()' と入力してください。 Eた、R や R のパッケージを出版物で引用する際の形式については citation()' と入力してください。		
demo() ' と入力すればデモをみることができます。 help() ' とすればオンラインヘルプが出ます。		
help.start()' で HTML ブラウザによるヘルプがみられます。 q()' と入力すれば R を終了します。		

うまく実行できると、以下のように表示されます。

	Ratouio			
- Mile - Addins -			Project: (None)	
• readData.R ×	Environment History Connections		-0	
🔅 🖉 📄 Source on Save 🔍 🎢 - 📄 👄 Run 🐤 🕞 Source - 🌊	😅 🕞 🖙 Import Dataset 👻 🅑		🗏 List + 🛛 🙆	
<pre>1 data <- read.csv("iris.csv", row.names = 1) 2 based(data)</pre>	Global Environment •	Q,		
3	Data			
	Gata 130 005. OF 5 Vertebles			
3:1 (Top Level) = R Script =	Files Plots Parkages Heln Viewer			
Console Terminal × Jobs ×	Q New Colder Q Dalata Rename A More -			
~/Desktop/Rtemp/W K は多くの夏献者による其同プロジェクトです。				
詳しくは 'contributors()' と入力してください。	A Name	Size	Modified	
<pre>citation()' と入力してください。 'demo()' と入力してください。 'help()' とすればオンラインヘルブが出ます。 'help.stort()' で HTML ブラウザによるヘルブがみられます。 'q()' と入力すれば R を続了します。</pre>	 .Rhistory iris.csv readData.R 	852 B 4.1 KB 55 B	Apr 9, 2020, 10:54 PM Apr 9, 2020, 10:14 PM Apr 9, 2020, 10:52 PM	
<pre>> data << read.csv("iris.csv", row.names = 1) > head(data) Sepal.Length Sepal.Width Petal.Length Petal.Width Species 1 5.1 3.5 1.4 0.2 setosa 2 4.9 3.0 1.4 0.2 setosa 3 4.7 3.2 1.3 0.2 setosa 4 4.6 3.1 1.5 0.2 setosa 5 5.0 3.6 1.4 0.2 setosa 6 5.4 3.9 1.7 0.4 setosa </pre>				

ここまでできれば、授業において躓くところは少ないでしょう。