Workshop on the Mathematical Modelling of Variant Replacement of Infectious Diseases Pathogens

RelRe: A command-line tool to predict the trajectory of variant replacement using relative reproduction numbers

Kimihito Ito
International Institute for Zoonosis Control
Hokkaido University

RelRe

- A command-line tool
  - to estimate relative reproduction numbers
  - to predict the trajectory of variant replacement
- Implemented in the Julia language
- Running on Windows, macOS, and Linux.
- Available on our GitHub page
  - https://github.com/KimihitoIto/RelRe

Installation of Julia (macOS)

- You can download the binaries MacOS at https://julialang.org/downloads/.
- Click on .dmg to download the installation package.
- Open the downloaded file and Drag and Drop the Julia-1.10.app to the Application folder to complete the installation.

Type the following command in Terminal or .bashrc (or .zshrc)

```
export PATH=/Applications/Julia-
1.10.app/Contents/Resources/julia/bin:$PATH
```

Installation of Julia (Linux)

- RelRe doesn’t work with the old Julia binary provided by the apt package manager in Ubuntu.
- Download the binaries Linux at https://julialang.org/downloads/.
- Copy the link to the binary to the clipboard

```
cd
mkdir Tools
cd Tools
```

Installation of Julia (Linux)

Type wget and paste the copied link to download the binary.

```
wget https://julialang-s3.julialang.org/bin/linux/x64/1.10/julia-1.10.5-linux-x86_64.tar.gz
```

Type the following to extract the contents.

```
tar xvfz julia-1.10.5-linux-x86_64.tar.gz
```

Type the following command in Terminal or .bashrc (or .zshrc)

```
export PATH=$HOME/Tools/julia-1.10.5/bin:$PATH
```

Installation of Julia (Windows)

- Open the Microsoft Store from the Start Menu of your windows.
- Type “ubuntu” in the search box. Install Ubuntu 24.04 LTS by clicking the “Get” button.
- Click on the Ubuntu icon, then a Terminal will appear.
- Set your login ID and password to complete the installation.
- Follow the Julia instructions for Linux in Section.

Run Julia to check installation

Type the following to run Julia.

```
julia
```

```
 _       _ _(_)_     |  Documentation: https://docs.julialang.org
  (_)     | (_) (_)    |
   _ _   _| |_  __ _   |  Type "?" for help, "]?" for Pkg help.
  | | | | | | |/ _` |  |
  | | |_| | | | (_| |  |  Version 1.8.5 (2023-01-08)
 _/ |\__'_|_|_|\__'_|  |  Official https://julialang.org/ release
|__/                   |

julia>
```

Type the following to exit from Julia.

```
julia
```

Installation of R

- The sample codes include R programs to visualize the estimation results.
- To run the R codes, install the R program on your computer if you don’t have it.
- Download the installer program of R from https://www.r-project.org/
- Run the downloaded installer.

Installation of Visual Studio Code

- Download Visual Studio Code from its homepage and install it on your computer. (https://code.visualstudio.com/)
- This is used to view and edit the program codes.
- You can use another text editor with which you are familiar.

Download the RelRe program

- The RelRe program is available at the following GitHub repository.
- https://github.com/KimihitoIto/RelRe

Use the following command to download the RelRe program.

```
cd ~/Desktop
git clone git@github.com:KimihitoIto/RelRe.git
```

Check whether you have the RelRe folder.

```
ls -l
drwxr-xr-x  5 itok itok 4096 Sep  9 15:02
RelRe
```

Check the Downloaded files

- install_packages.jl (a script for installation)
- RelRe.jl (the RelRe program)
- README.md (general information)
- sample_data/ (a folder of sample datasets)

```
cd RelRe/
ls
```

```
-rw-rw-r-- 1 itok itok    89 Sep  9 15:02 install_packages.jl
-rw-r--r-- 1 itok itok  4771 Aug 22 17:13 README.md
-rw-rwxr-- 1 itok itok 22444 Aug 22 17:31 RelRe.jl
drwxr-xr-x 4 itok itok  4096 Sep  5 19:17 sample_data
```

Install Required Julia packages

- The RelRe program uses CSV, Dates, DataFrames, Distributions, NLopt, and ArgParse packages, all included in Julia’s General registry.
- Type the following command to install the required packages

```
julia install_packages.jl
```

It will take about 5-10 minutes to complete.

Sample Datasets

- The RelRe program contains sample datasets.
  - SARS-CoV-2-BA2-Tokyo-weekly
  - SARS-CoV-2-GT-Denmark

Type the following commands to check.

```
cd sample_data
ls -l
```

```
drwxr-xr-x 2 itok … Sep  5 19:19 SARS-CoV-2-BA2-Tokyo-weekly
drwxr-xr-x 2 itok … Sep  9 14:45 SARS-CoV-2-GT-Denmark
```

SARS-CoV-2-BA2-Tokyo-weekly

- Sample data and scripts to analyze weekly counts of Omiron BA.1 and BA.2 in Tokyo.

Type the following commands to check files.

```
cd SARS-CoV-2-BA2-Tokyo-weekly
ls -l
```

```
-rw-r--r-- 1 itok … Aug  9 09:53 Makefile
-rw-r--r-- 1 itok … Aug  9 09:53 plot_average_k.R
-rw-r--r-- 1 itok … Aug  9 09:53 plot_frequency_wkly.R
lrwxrwxrwx 1 itok … Aug  9 09:53 RelRe.jl -> ../../RelRe.jl
-rw-r--r-- 1 itok … Aug  9 09:53 Tokyo_BA1_BA2.csv
```

Input File

- RelRe reads the count data of variants from a CSV file
- Dates are assumed to be given in the YYYY-MM-DD format

date_from	**date_till**	**number of *a***	**number of A_1**	**…**	**number of A_n**
⋮	⋮	⋮	⋮	⋮	⋮
t_h	u_h	$N_a(o_h)$	$N_{A_1}(o_h)$	$\cdots$	$N_{A_1}(o_h)$
⋮	⋮	⋮	⋮	⋮	⋮

Tokyo_BA1_BA2.csv

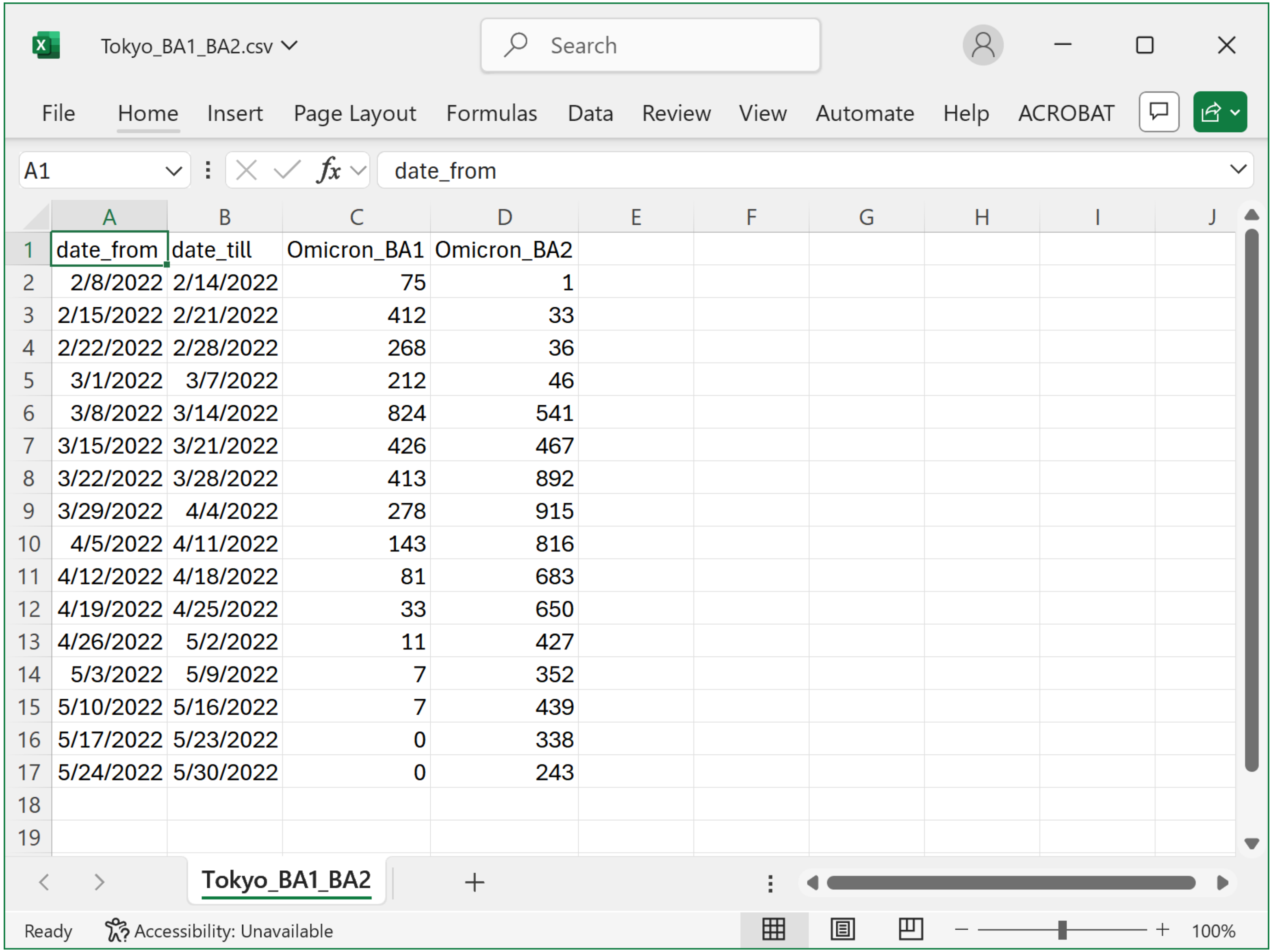
- Weekly count data of Omiron BA.1 and BA.2 in Tokyo.

Type the following commands to check the data.

```
cat Tokyo_BA1_BA2.csv
```

```
date_from,date_till,Omicron_BA1,Omicron_BA2
2022-02-08,2022-02-14,75,1
2022-02-15,2022-02-21,412,33
2022-02-22,2022-02-28,268,36
2022-03-01,2022-03-07,212,46
2022-03-08,2022-03-14,824,541
2022-03-15,2022-03-21,426,467
2022-03-22,2022-03-28,413,892
…
```

Tokyo_BA1_BA2.csv



date_from	date_till	Omicron_BA1	Omicron_BA2
2/8/2022	2/14/2022	75	1
2/15/2022	2/21/2022	412	33
2/22/2022	2/28/2022	268	36
3/1/2022	3/7/2022	212	46
3/8/2022	3/14/2022	824	541
3/15/2022	3/21/2022	426	467
3/22/2022	3/28/2022	413	892
3/29/2022	4/4/2022	278	915
4/5/2022	4/11/2022	143	816
4/12/2022	4/18/2022	81	683
4/19/2022	4/25/2022	33	650
4/26/2022	5/2/2022	11	427
5/3/2022	5/9/2022	7	352
5/10/2022	5/16/2022	7	439
5/17/2022	5/23/2022	0	338
5/24/2022	5/30/2022	0	243

Command Example (1)

```
julia RelRe.jl -b Omicron_BA1 -a 4.03 -t
0.735 -i Tokyo_BA1_BA2.csv
```

-b the name of the baseline variant

Use Omicron_BA1 as the baseline variant

Command Example (1)

```
julia RelRe.jl -b Omicron_BA1 -a 4.03
 -t 0.735 -i Tokyo_BA1_BA2.csv
```

-a shape parameter

Give the shape parameter of gamma distribution for generation time

-t scale parameter

Give the scale parameter of gamma distribution for generation time

Command Example (1)

```
julia RelRe.jl -b Omicron_BA1 -a 4.03 -
t 0.735 -i Tokyo_BA1_BA2.csv
```

-i filename

Give the input file containing temporal count data of variants

Command Example (1)

Type the following commands to run the analysis.

```
julia RelRe.jl -b Omicron_BA1 -a 4.03 -t
0.735 -i Tokyo_BA1_BA2.csv
```

```
parsed_args = Dict{String, Any}("estimate_GT" => false,
"estimate_CI" => false, "ftol_abs" => 0.0, "theta" => 0.735,
"maxeval" => 5000000, "out" => "", "len" => -1, "in" =>
"Tokyo_BA1_BA2.csv", "start" => "", "subjects" => Symbol[],
"alpha" => 4.03, "end" => "", "ftol_rel" => 1.0e-8, "future" => 0,
"Dirichlet" => false, "division" => 1, "frequency" => false,
"baseline" => :Omicron_BA1, "undetected" => false)
Loading counts
Baseline
Omicron_BA1
…
```

It takes less than 1 minute to finish the analysis.
A message of “done” is shown when it finishes.

Output files

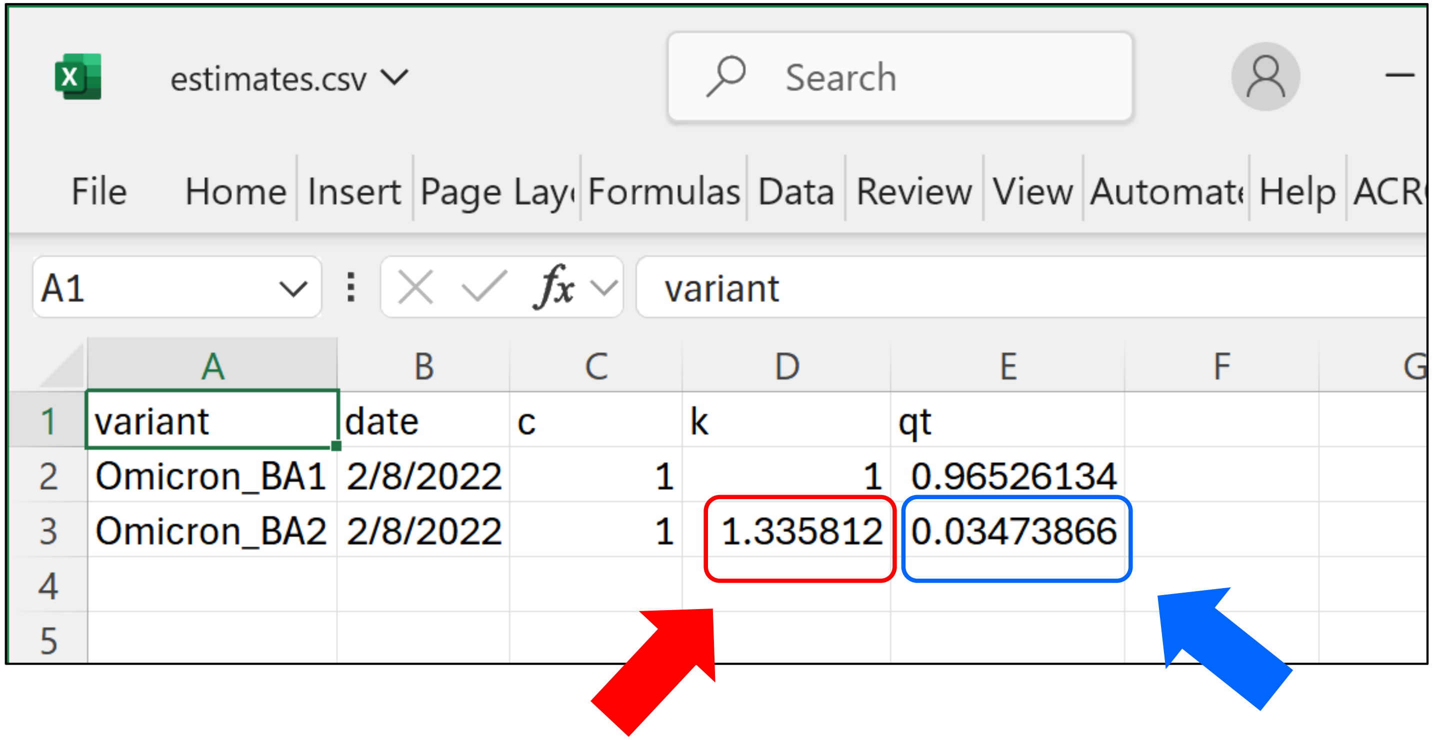
- RelRe outputs the following files
  - estimates.csv
  - loglikelihood.csv

Type the following command to check.

```
ls -l
```

```
-rw-rw-r-- 1 itok itok  137 Sep  9 18:10 estimates.csv
-rw-rw-r-- 1 itok itok   59 Sep  9 18:10 loglikelihood.csv…
```

estimates.csv



variant	date	c	k	qt
Omicron_BA1	2/8/2022	1	1	0.96526134
Omicron_BA2	2/8/2022	1	1.335812	0.03473866

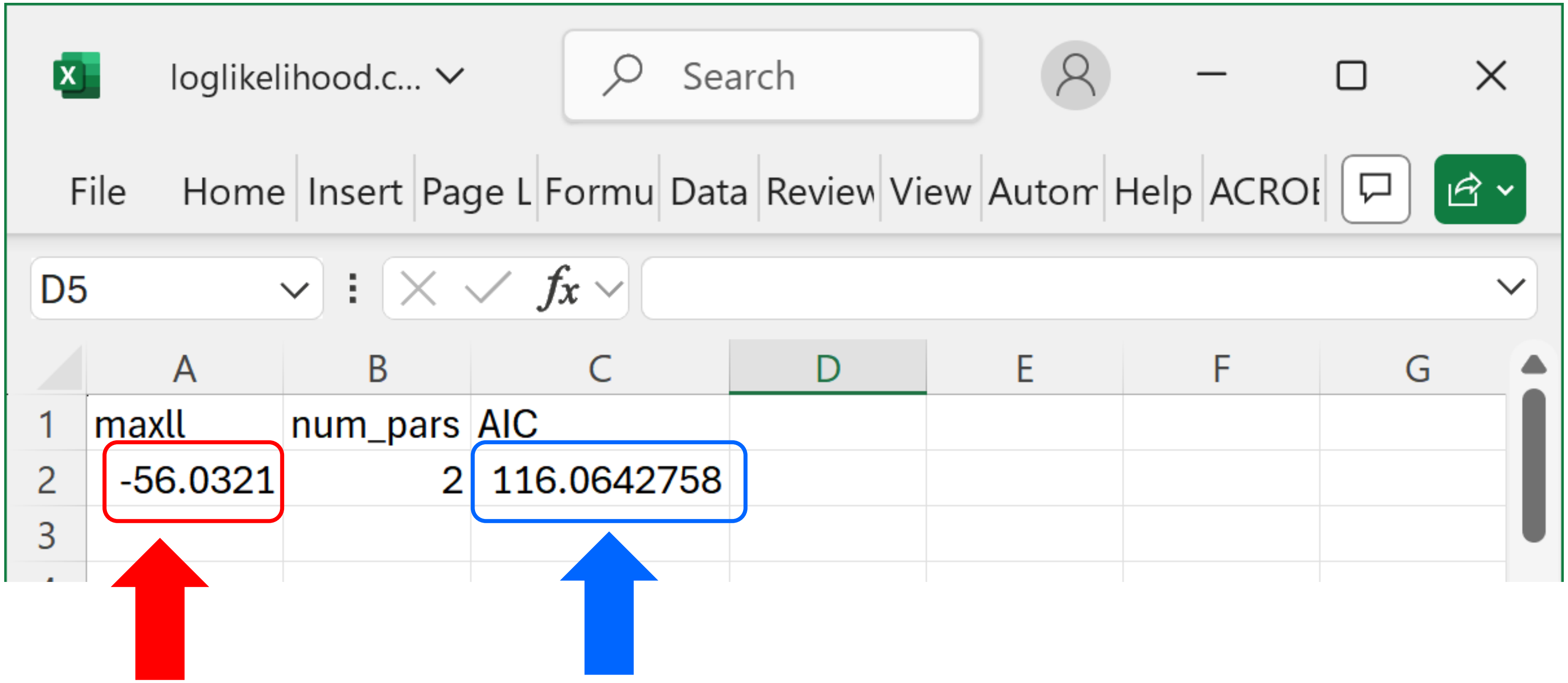
Relative reproduction number of Omicron_BA2 w.r.t. Omicron_BA1. BA.2 is 1.34 times more transmissible than BA.1.

The initial frequency of Omicron_BA2

Akaike Information Criterion

- Akaike Information Criterion (AIC) of a model is defined as
  $$AIC = 2k - 2\log\left(\hat{L}\right),$$
  where k is the number of the parameters of the model and $\hat{L}$ is the maximum log-likelihood of model parameters (Akaike, 1975).
- For a given set of observations, a model with a smaller AIC is considered a better model than a larger one.

loglikelihood.csv



	A	B	C
1	maxll	num_pars	AIC
2	-56.0321	2	116.0642758
3			

The maximum likelihood

Akaike Information Criterion (AIC) of the model

Command Example (2)

```
julia --threads 4 RelRe.jl -b Omicron_BA1 -
a 4.03 -t 0.735 -i Tokyo_BA1_BA2.csv -c
```

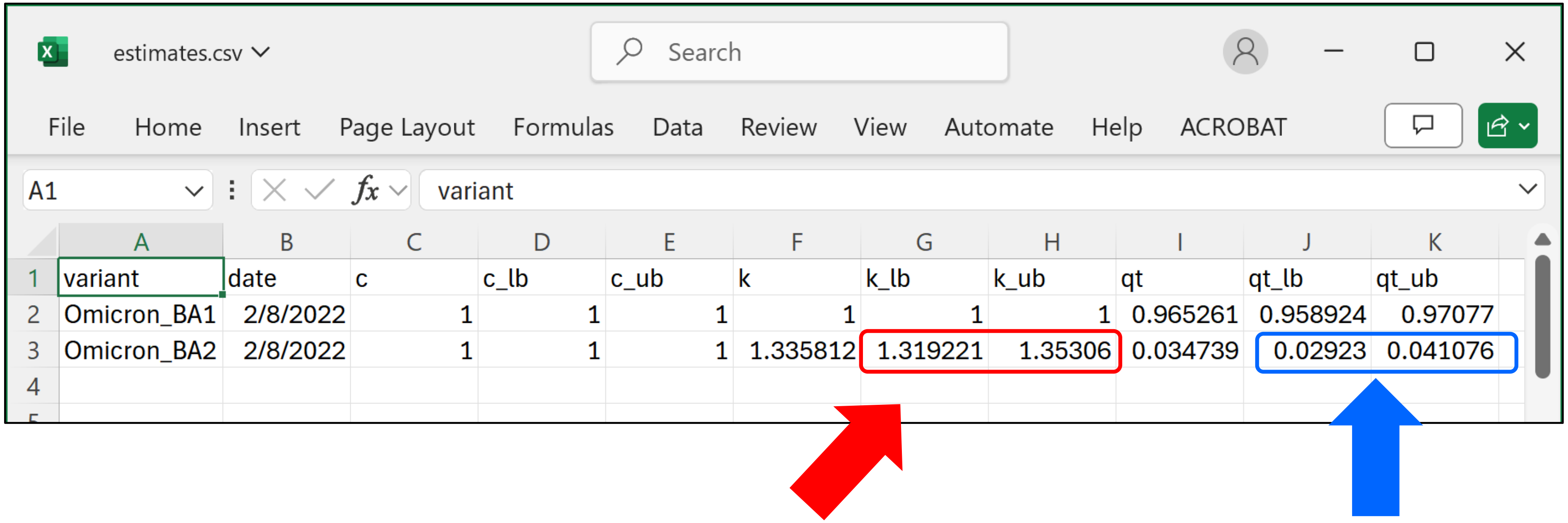
-c

Use -c option for calculating 95% confidence intervals of the estimates.

--threads N

Use Julia’s –threads option for using multiple threads.

estimates.csv



	A	B	C	D	E	F	G	H	I	J	K
1	variant	date	c	c_lb	c_ub	k	k_lb	k_ub	qt	qt_lb	qt_ub
2	Omicron_BA1	2/8/2022	1	1	1	1	1	1	0.965261	0.958924	0.97077
3	Omicron_BA2	2/8/2022	1	1	1	1.335812	1.319221	1.35306	0.034739	0.02923	0.041076
4											

The 95% confidence intervals of relative reproduction number of Omicron_BA2 w.r.t. Omicron_BA1.

The 95% confidence intervals of the initial frequency of Omicron_BA2

Command Example (3)

```
julia --threads 4 RelRe.jl -b Omicron_BA1 -
a 4.03 -t 0.735 -i Tokyo_BA1_BA2.csv -c -q
-f 60
```

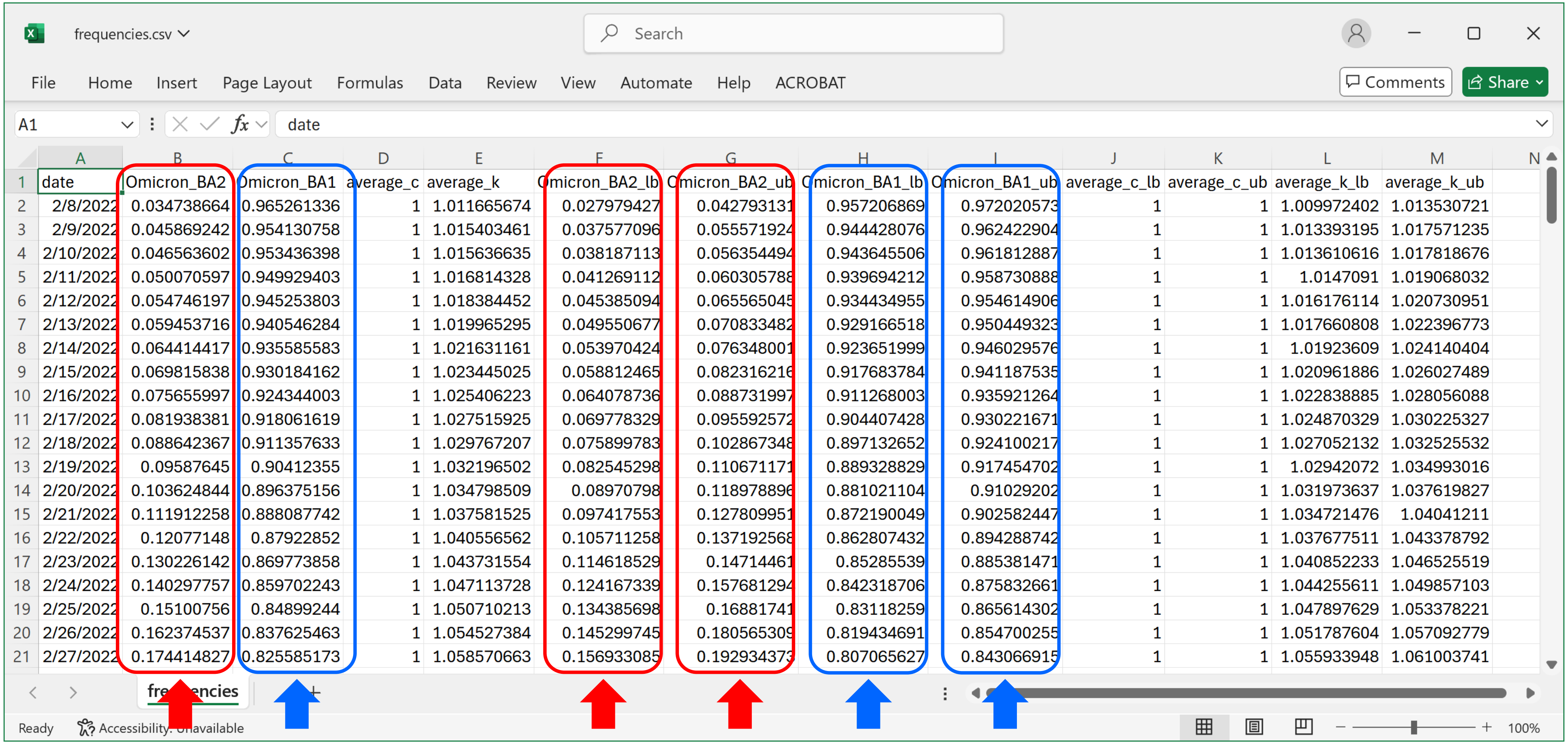
-q

Use -q option to output the relative frequencies of variants estimated by the model.

-f N (days)

Use –q option to specify the duration of prediction in the future

frequencies.csv



	A	B	C	D	E	F	G	H	I	J	K	L	M
1	date	Omicron_BA2	Omicron_BA1	average_c	average_k	Omicron_BA2_lb	Omicron_BA2_ub	Omicron_BA1_lb	Omicron_BA1_ub	average_c_lb	average_c_ub	average_k_lb	average_k_ub
2	2/8/2022	0.034738664	0.965261336	1	1.011665674	0.027979427	0.042793131	0.957206869	0.972020573	1	1	1.009972402	1.013530721
3	2/9/2022	0.045869242	0.954130758	1	1.015403461	0.037577096	0.055571924	0.944428076	0.962422904	1	1	1.013393195	1.017571235
4	2/10/2022	0.046563602	0.953436398	1	1.015636635	0.038187113	0.056354494	0.943645506	0.961812887	1	1	1.013610616	1.017818676
5	2/11/2022	0.050070597	0.949929403	1	1.016814328	0.041269112	0.060305788	0.939694212	0.958730888	1	1	1.0147091	1.019068032
6	2/12/2022	0.054746197	0.945253803	1	1.018384452	0.045385094	0.065565045	0.934434955	0.954614906	1	1	1.016176114	1.020730951
7	2/13/2022	0.059453716	0.940546284	1	1.019965295	0.049550677	0.070833482	0.929166518	0.950449323	1	1	1.017660808	1.022396773
8	2/14/2022	0.064414417	0.935585583	1	1.021631161	0.053970424	0.076348001	0.923651999	0.946029576	1	1	1.01923609	1.024140404
9	2/15/2022	0.069815838	0.930184162	1	1.023445025	0.058812465	0.082316216	0.917683784	0.941187535	1	1	1.020961886	1.026027489
10	2/16/2022	0.075655997	0.924344003	1	1.025406223	0.064078736	0.088731997	0.911268003	0.935921264	1	1	1.022838885	1.028056088
11	2/17/2022	0.081938381	0.918061619	1	1.027515925	0.069778329	0.095592572	0.904407428	0.930221671	1	1	1.024870329	1.030225327
12	2/18/2022	0.088642367	0.911357633	1	1.029767207	0.075899783	0.102867348	0.897132652	0.924100217	1	1	1.027052132	1.032525532
13	2/19/2022	0.09587645	0.90412355	1	1.032196502	0.082545298	0.110671171	0.889328829	0.917454702	1	1	1.02942072	1.034993016
14	2/20/2022	0.103624844	0.896375156	1	1.034798509	0.08970798	0.118978896	0.881021104	0.91029202	1	1	1.031973637	1.037619827
15	2/21/2022	0.111912258	0.888087742	1	1.037581525	0.097417553	0.127809951	0.872190049	0.902582447	1	1	1.034721476	1.04041211
16	2/22/2022	0.12077148	0.87922852	1	1.040556562	0.105711258	0.137192568	0.862807432	0.894288742	1	1	1.037677511	1.043378792
17	2/23/2022	0.130226142	0.869773858	1	1.043731554	0.114618529	0.14714461	0.85285539	0.885381471	1	1	1.040852233	1.046525519
18	2/24/2022	0.140297757	0.859702243	1	1.047113728	0.124167339	0.157681294	0.842318706	0.875832661	1	1	1.044255611	1.049857103
19	2/25/2022	0.15100756	0.84899244	1	1.050710213	0.134385698	0.16881741	0.83118259	0.865614302	1	1	1.047897629	1.053378221
20	2/26/2022	0.162374537	0.837625463	1	1.054527384	0.145299745	0.180565309	0.819434691	0.854700255	1	1	1.051787604	1.057092779
21	2/27/2022	0.174414827	0.825585173	1	1.058570663	0.156933085	0.192934373	0.807065627	0.843066915	1	1	1.055933948	1.061003741

The relative frequencies of BA2 and BA1 with their 95% confidence intervals.

R scripts for plotting RelRe outputs

- RelRe is accompanied with R scripts for plotting its outputs

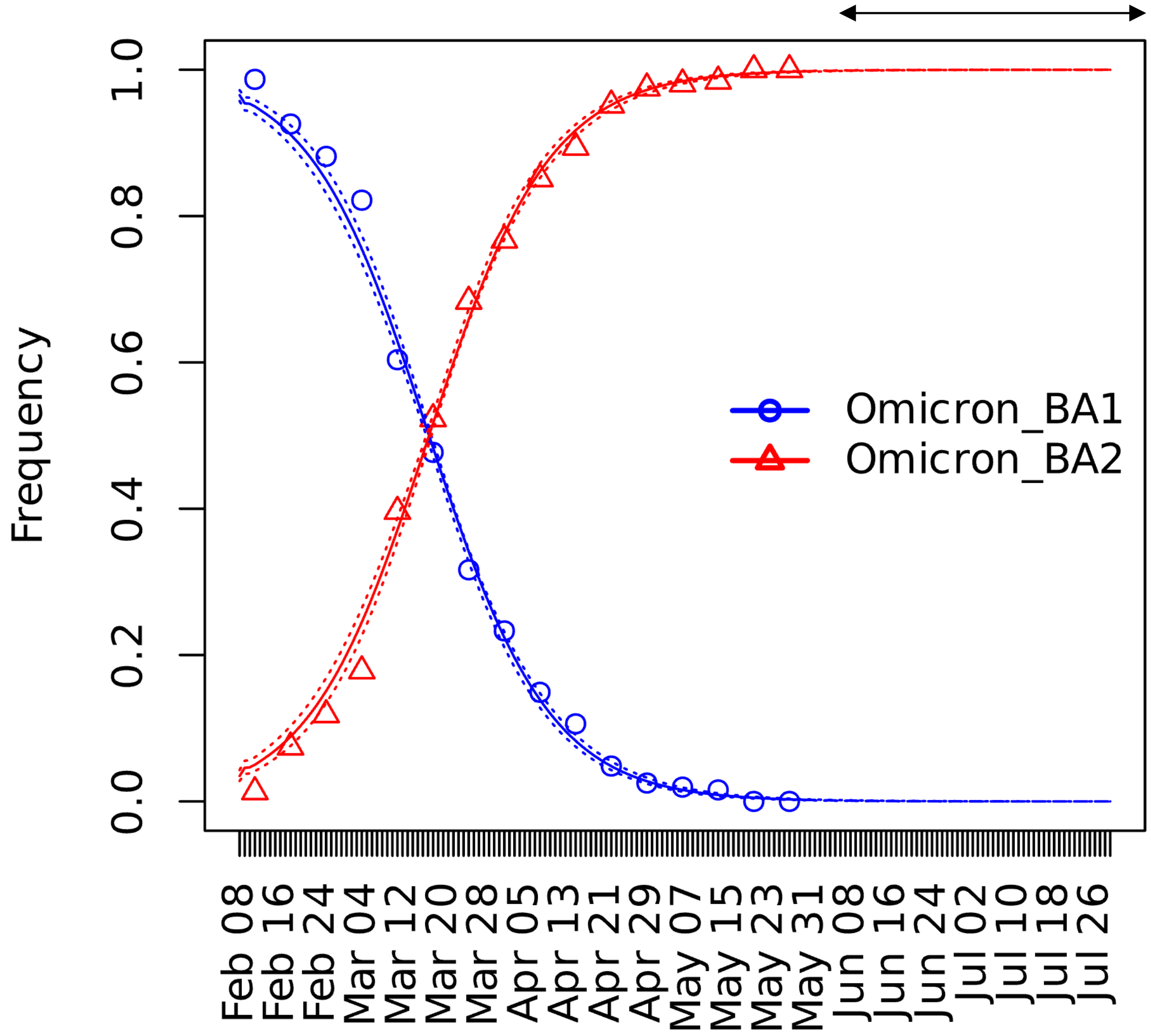
Type the following command to create a pdf file

```
R -f plot_frequency_wkly.R
```

The R script uses the dplyr package of R. Install it by typing the following command in R if you don't have it.

```
R
> install.packages("dplyr")
> q()
```

frequency.pdf



Predicted duration
specified by -f option

Makefile

- The sample programs contain Makefile to run analyses.
- Just type “make” in the command line to run the program
- Type “make clean” to remove the result files.

```
make
```

```
make clean
```

parameter	description
-a, --alpha	shape parameter of gamma distribution for generation time (type: Float64, default: 2.03)
-b, --baseline	the variant used as the baseline of relative reproduction numbers
-c, --estimate_CI	estimate 95% confidence intervals
-d, --division	the number of time steps for each day (type: Int64, default: 1)
-D, --Dirichlet	use Dirichlet multinomial as the observation model
-e, --end	end date of the analysis (default: "")
-f, --future	duration in days for predicting variant frequencies (type: Int64, default: 0)
-g, --estimate_GT	estimate relative generation times of variants
-h, --help	show the help message and exit
-i, --in	input file containing temporal count data of variants
-j, --subjects	list of variants to calculate relative reproduction numbers (type: Symbol)
-l, --len	truncation point of gamma distribution for generation time (type: Int64, default: automatic)
-u, --undetected	assume all variants exist undetected from the start date
-o, --out	prefix of output files (default: "")
--ftol_abs	stopping criterion used as ftol_abs in NLopt (type: Float64, default: 0.0)
--ftol_rel	stopping criterion used as ftol_rel in NLopt (type: Float64, default: 1.0e-8)
--maxeval	stopping criterion used as maxeval in NLopt (type: Int64, default: 5,000,000)
-q, --frequency	calculate the time course of variant frequencies
-s, --start	start date of the analysis (default: "")
-t, --theta	scale parameter of gamma distribution for generation time (type: Float64, default: 2.32)

Summary

- The RelRe program estimates relative reproduction numbers among variants from their count data.
- The RelRe needs parameters of the generation time distribution.
- The RelRe program can predict the trajectory of variant replacement.