M19

Use this package to generate cut-offs using the maximal rank statistic method.

Installation

For generic installation and set-up of R Studio, please see guides here and here.

First install and load the 'devtools' package;

```
install.packages("devtools")
library(devtools)
```


Figure 1: Console - When the vertical dash is flickering, the installation and loading is finished.

Next, you need to install the 'gplots' package;

```
install.packages("gplots")
```


Figure 2: Console - When the vertical dash is flickering, the installation and loading is finished.

Finally, you need to install the 'M19' package;

```
install_github("HHayman/M19")
```

```
> library(devtools)
Loading required package: usethis
>
>
>
> install_github("HHayman/M19")
Downloading GitHub repo HHayman/M19@HEAD
√  checking for file 'C:\Users\liamj\AppData\Local\Temp\RtmpGoB1yC\remotes10882de24d3f\HHayman-M19-436bbbf/DESCRIPTION' ...
-  preparing 'M19':
√  checking DESCRIPTION meta-information ...
-  checking for LF line-endings in source and make files and shell scripts
-  checking for empty or unneeded directories
-  building 'M19_0.1.0.tar.gz'

Installing package into 'C:/Users/liamj/Documents/R/win-library/4.0'
(as 'lib' is unspecified)
* installing *source* package 'M19' ...
** using staged installation
** R
** byte-compile and prepare package for lazy loading
** help
*** installing help indices
  converting help for package 'M19'
    finding HTML links ... done
    ImmSurv                                 html
    rCut                                    html
** building package indices
** testing if installed package can be loaded from temporary location
*** arch - i386
*** arch - x64
** testing if installed package can be loaded from final location
*** arch - i386
*** arch - x64
** testing if installed package keeps a record of temporary installation path
* DONE (M19)
> library(M19)
> |
```

Figure 3: Console - When the vertical dash is flickering, the installation and loading is finished.

Preparing Data

You need to carry out any exclusions etc in your data set - include the minimum amount of data possible. Your final data should include columns such as (example); an identifier (TMA_ID), survival time(s) (CSS_Time), corresponding survival status(s) (CSS_status), your variables (CD8_Tumour, CD8_Healthy).

	A	B	C	D	E	F	G
1	TMA_ID	a0cd1ncd2_2017	GD_Tumour_Stroma_Perc	GD_Tumour_Epithelium_Perc	GD_Healthy_LaminaPropria_Perc	GD_Healthy_Epithelium_Perc	CSS_2017

Coding is based on those used in the Glasgow combined cohort, if your variables don't match, use the 'MISC' option or let me know.

CSS/OS should be coded as 0 = alive, 1 = CD, 2 = NCD. The function will split it into CSS and OS for you. DFS considers values >0 to be an event. RFS considers 1 to be a recurrence event. MISC considers 1 a miscellaneous event, everything else becomes 0. You can use this with anything, but code your data appropriately.

First, you need to read your data file. It should be a CSV file with no spaces in the column names, use '_'. To import your data, use this code and use the pop up window to select your data file;

```
Data <- read.csv(file.choose(), fileEncoding = 'UTF-8-BOM')
```

Produce Cut-Offs

Check your data file in R and if you're happy with it, edit and use the code below to call the function;

All variable names must match exactly. You'll get an error if they don't.

For CSS, OS, DFS, RFS and MISC, Use 'Yes' if you want that cut-off to be generated, and 'No' if you do not. It is case sensitive and an input is required. If you state 'No', you don't need to provide any other variables for that outcome.

For 'Survival', 'SurvivalStatus' and any other required time and status variables, replace the text within the quotation marks with your variable names.

minprop sets a minimum proportion of cases to be included in any group. The default is 0.1, which is standard.

PlotPalette determines the colour scheme. "SPSS" uses the colour scheme from SPSS, "Grayscale" is pseudo black and white.

Replace 'Variable1' etc with your variables. There is no limit so just add/remove arguments as appropriate.

If anything is unclear, ask, it's much better than having to redo your analysis :)

```
rCut(Data, CSS = "Yes", OS = "Yes", DFS = "Yes", RFS = "Yes", MISC = "No", minprop = 0.1, PlotPalette =
```

When you get a message from Wonder Woman and the flashing vertical dash is back, the function has finished running. It might take a few seconds.

```
> rCut(Data, CSS = "Yes", OS = "No", PFS = "No", PlotPalette = "SPSS", ID = "ID", Survival = "CSS_2017", SurvivalStatus = "a0cd1ncd2_2017", variables =
 c("GD_Tumour_Stroma_Perc", "GD_Tumour_Epithelium_Perc", "GD_Healthy_LaminaPropria_Perc", "GD_Healthy_Epithelium_Perc"))
Wonder Woman hopes that you enjoy your cut-offs. Blessings of Gaea be with you :)
> |
```

Figure 4: Console - When the vertical dash is flickering, the installation and loading is finished.

I suggest ignoring the output within R studio itself, but the bottom right pane of R studio will contain a copy of your plots. If you want to look at them within R studio, use the arrow on the left of the ribbon to cycle through the plots.

If you open your working directory folder you will find a folder named 'CutOffs_Data_#'. Every time you run the function you get a fresh folder in your working directory, numbered incrementally if a folder already exists.

Within the folder you will find your data; * A pdf file named 'YourPlots' containing a text print of your cut offs and associated test statistic, and all of your plots * Individual png copies of each plot for use in presentations etc. These are labeled based on your variables and event outcomes * Two csv files, 'OriginalData' and 'ModifiedData'. These are copies of your original data and the modified version from the function. Leave them alone, they are just for future reference if you need to know exactly what data you used * A txt file called 'Waste'. This is just a dump for anything that we don't send to the console, thus avoiding clutter. Ignore it
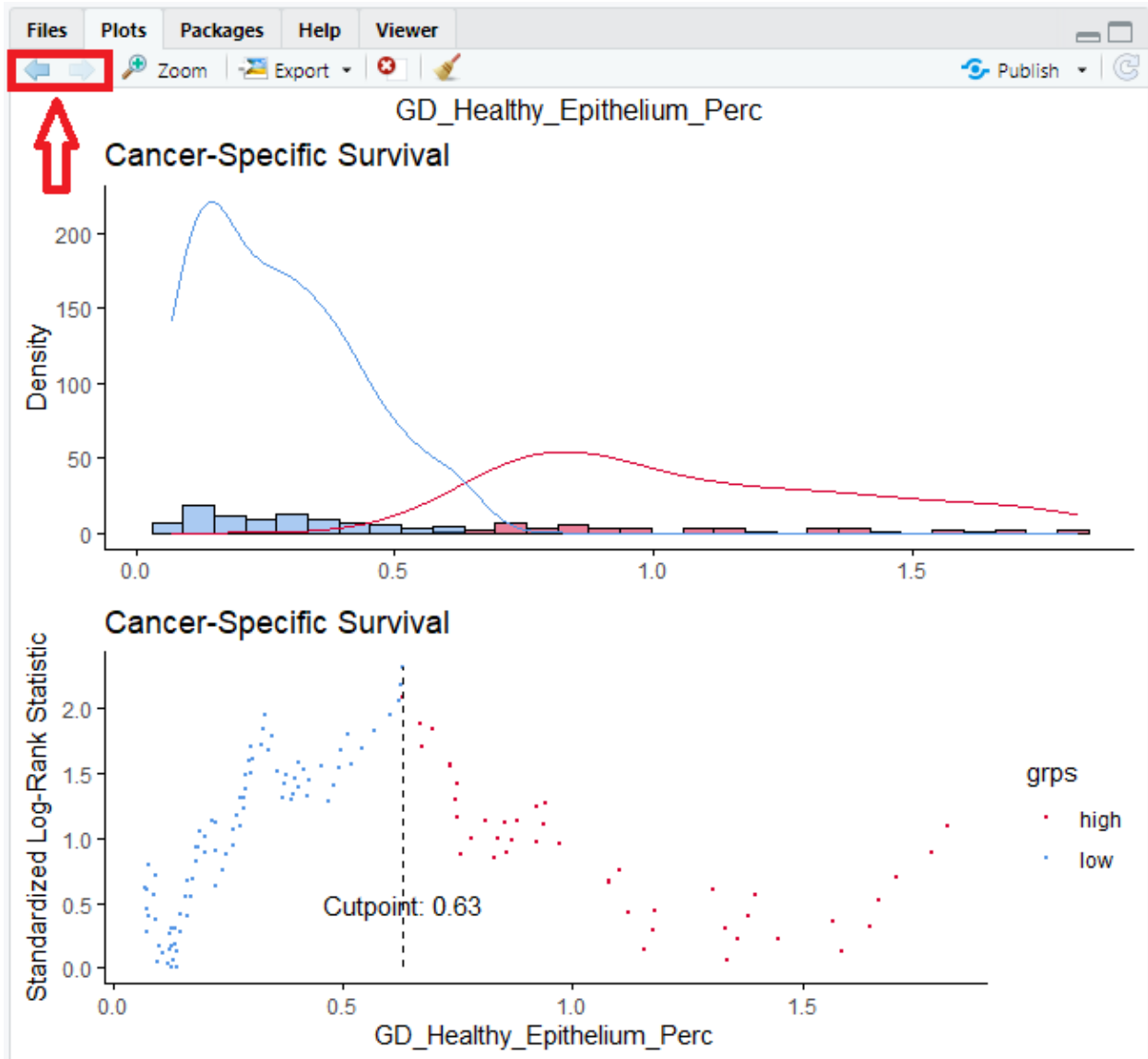
Figure 5: Arrows are highlighted by the red arrow and box.

Figure 6: Your data folder within your R working directory.

Name	Date modified	Type	Size
CSS_GD_Healthy_Epithelium_Perc	05/01/2021 14:16	PNG File	7 KB
CSS_GD_Healthy_LaminaPropria_Perc	05/01/2021 14:16	PNG File	8 KB
CSS_GD_Tumour_Epithelium_Perc	05/01/2021 14:16	PNG File	7 KB
CSS_GD_Tumour_Stroma_Perc	05/01/2021 14:16	PNG File	8 KB
ModifiedData	05/01/2021 14:16	CSV File	12 KB
OriginalData	05/01/2021 14:16	CSV File	11 KB
waste	05/01/2021 14:16	Text Document	1 KB
YourPlots	05/01/2021 14:16	Microsoft Edge P...	72 KB

Figure 7: Your data folder within your R working directory.