Overview of resources for

A Biostatistics Toolbox for Data Analysis,
by Steve Selvin, Cambridge University Press, 2015

The resources at the website **www.cambridge.org/9781107113084** consist of four elements labeled:

- R-lessons,
- R-graphics,
- R-toolbox and
- R-case.studies.

These resources support the use of the R statistical language for the data analysis methods presented in the book. The R language is free and open source. For more information about R, and to download the software, see the R home page: **www.r-project.org/**.

R-lessons (9 files)

The R-lessons folder holds nine groups of R-commands labeled *lessons.1.r*, *lessons.2.r*, ..., *lessons.9.r*. The "lessons" contain an extensive number of examples of R-commands from simple and often used to less simple and less often used. Each lesson is a series of illustrations generated by R-commands typically used in statistical data analyses.

For details about each command, such as usage, options, methods, related commands and references, use R’s built-in help facility: all R-commands preceded by a question mark produce a complete description of the command. For example, the command *?cbind* creates a complete and very detailed description of the command *cbind* used to create vectors and matrices. The nine lessons are intended as a rich source of illustrations rather than documentation.

R-graphics (12 files)

The files in the R-graphics folder play a role similar to the files in R-lessons. The twelve R-coded plot routines give a variety of graphic examples. These illustrations serve as a starting point to understanding the process of using R-commands to generate plots and graphs in connection with analysis of data. An important command is *expression*, which allows the creation of graphic symbols and features not available from a computer keyboard. For example, the R-code *expression(e^{-t*lambda})* yields the graphic expression $e^{-t\lambda}$. Two pages of complete descriptions and options are available, as usual, using the R-command *?plotmath*.

R-toolbox (24 files)

The R-toolbox folder contains a large number of R-generated analytic software. The same files are contained in the appendix of the text *A Biostatistical Toolbox for Data Analysis*. The files can be downloaded and used among many possibilities, to verify the examples in the text, identify the influences of outlier observations and sample size. Of course, different sets of data can be explored with these analytic methods. In addition, the details of R-code that created the computer analysis often provides valuable insight into the details of the statistical method itself.

R-case.studies (21 files)

The files contained in the R-case.studies folder follow the pattern of the R-coded methods in the R-toolbox folder. However, these examples use complete and real data, providing a comprehensive description of the case studies presented in the text. Eleven of these analyses require large data sets that are also included in the folder. For example, an R-program labeled *chd.r* requires the corresponding data file labeled *chd.data* that contains seven variables and 3153 observations.

To link the analysis R-code to the data file, the R-command *read.table* is used. It is necessary to modify the pathname to correspond to the specific computer to be used. An example of a specific command that reads the data contained in the file *chd.data* is found in the file *chd.r*. Supplying the data file path-name (contained between the quotation marks) creates the link between code and data.