Tianran Zhang, Wenyu Zhu, Yinuo Liu

R package

Feb 23, 2020

Design and Analysis of Biomedical Studies

Tutorial

The tutorial is for the SCR package, an educational-based R package, providing the references of exercises from the book Statistical Computing with R by Maria L. Rizzo.

In many cases, when students are doing coding exercises, it is inconvenient for them to check whether the results are correct, and inefficient to find typos that cause errors. To help improve the efficiency of students' self-learning as well as for professors correcting the coding homework, we decided to design an R package for educational purposes that anyone who uses R or Stats for work or study can benefit from.

After some screening, we chose Statistical Computing with R by Maria L. Rizzo based on two principles. First, Statistical Computing with R is a famous fundamental textbook that is not only used widely in many basic Statistical courses but also referred to by many Statistical Computing self-learners. Besides, there is no official solution manuscript of the exercises for this book that people who use it can refer to. Thus, we built this package related to this book.

The current work contained seven chapters of the book from chapter three to chapter nine. Chapter three mainly discusses several kinds of methods to simulate random variables from specified probability distributions. Chapter four explores different methods for visualizing multivariate data. Chapter five introduces one of the most popular random sampling methods—Monte Carlo iteration. Chapter six further explores the Monte Carlo method in inference, which contains the estimation and the hypothesis test. Chapter seven discusses several kinds of bootstraps. Chapter eight mainly explores the permutation tests by the bootstrap simulations. Chapter nine discusses the Markov Chain Monte Carlo (MCMC) methods mainly by four methods.

We divided those chapters into 3 parts and each of us was responsible for one of this part. Then we made functions of the exercises and wrote R documents of them. Having finished some

exercises, we found that there is some general knowledge in these exercises, which can be expanded to help learners better understand the topics in Rizzo's book. Therefore, we decided to make expanded functions to help students better understand the topics in Rizzo's book, furthermore, to help professors better present the intrinsic meanings of the problems.

So far, we made two expanded functions in the package. One is MCsimCL1, which is for the Monte Carlo simulation to estimate confidence level. The other is MCMCsim9, which is visualizing the differences in the Markov Chain generated by these three methods. To make functions more readable and usable to our potential package users, we spent much time polishing the R documentation, trying to display the ideas of solving each question and specify the meanings of each parameter in our functions.

There are more than 100 exercises in this book, ranging from some basic statistical problems to higher level simulation methods. Up till now, we have finished a total of 75 functions, containing 73 exercises in Rizzo's book and 2 extended functions in our R-package. Finishing this package is quite a challenge for us and needs huge endeavor, but we do enjoy the process that three of us all spare no effort contributing to this package. We will try our best to finish it in the nearly future and publish it online and look forward to seeing it can truly benefit others who are struggling with statistical computing.