STAT 513/413: Lecture 1
Orientation

(the choices made)

Welcome to...

STAT 513: Computational Statistics

Assignment Project Exam Help

https://powcoder.com

Add WeChat powcoder

Welcome to...

STAT 513: ~~Computational Statistics~~ Statistical Computing

and

STAT 413: Computational Statistics

Welcome to...

STAT 513: ~~Computational Statistics~~ Statistical Computing

and

STAT 413: ~~Computational Statistics~~

Introduction to Computing for Data Science

Assignment Project Exam Help

https://powcoder.com

Add WeChat powcoder

Welcome to...

STAT 513: ~~Computational Statistics~~ Statistical Computing

and

STAT 413: ~~Computational Statistics~~

Introduction to Computing for Data Science

Two courses in one room...

1

Welcome to...

STAT 513: ~~Computational Statistics~~ Statistical Computing

and

STAT 413: ~~Computational Statistics~~
Introduction to Computing for Data Science

Two courses in one room...

...online... Once again:

Welcome to...

STAT 513: Computational Statistics   Statistical Computing

and

STAT 413: Computational Statistics

Introduction to Computing for Data Science

Assignment Project Exam Help

https://powcoder.com

Add WeChat powcoder

Two courses in one room...

...online...

Words, words, words!

Statistical Computing: what it is *not*

Rizzo has it pretty much right

Computational Statistics: may suggest to some minds (despite the calendar description) that it is all about obtaining the results of various statistical methods - regression, ANOVA, etc. - in some computing environment (with packaged software)

THIS IS NOT WHAT THIS COURSE INTENDS TO DO

This is rather what is done in topical courses in statistics, like those covering sample survey (STAT 361), design and analysis of experiments (STAT 368), regression (STAT 378), survival analysis (STAT 432), time series (STAT 479), ...

So what is STAT 513 then?

STATISTICAL ← COMPUTING

algorithmic and computational components useful in implementing statistical and similar methods

(with special emphasis on those that are specific for the field, *not just general-purpose numerical methods*)

and also

STATISTICAL → COMPUTING

algorithmic strategies inspired by statistical and probabilistic ideas

Nothing else than in the Calendar

STAT 513. Introduction to contemporary computational culture: reproducible coding, literate programming. Monte Carlo methods: random number generation, variance reduction, numerical integration, statistical simulations. Optimization (linear search, gradient descent, Newton-Raphson, method of scoring, and their specifics in the statistical context), EM algorithm. Fundamentals of convex optimization with constraints. Prerequisites: consent of the instructor.

A grad course aimed at grad students in statistics and related fields. Something like STAT 512 (Methods of Mathematics for Statistics) is for them regarding mathematics, STAT 513 (note the number) should be in computing

STAT 413: A spin-off for honors Stat

...but eventually acquiring its own life...

STAT 413. Survey of contemporary languages/environments suitable for algorithms of Statistics and Data Science. Introduction to Monte Carlo methods, random number generation and numerical integration in statistical context and optimization for both smooth and constrained alternatives, tailored to specific applications in statistics and machine learning. Prerequisites: STAT 265 or consent of the instructor.

...if not for budget cuts and contagious viruses and other disasters

This description may need some further evolution: it may include some of the statistical techniques that are not covered by other topical corses - but still by Rizzo's book - and on the other hand, it may drop some advanced parts - not covered by Rizzo's book anyway

For now, we all have to put up with combined STAT 513/413

STAT 513 is more advanced - a grad course - so it assumes more

In particular parts indicated as "STAT 413" are supposed to be easy to learn, if not already known, by STAT 513 students (so they are compulsory for them too)

On the other hand, parts indicated as "STAT 513" may be...

... oh no, not omitted, definitely not - but, say, expected to be graded less rigorously (say, homework yes, but final not)

Also: course objectives and expected learning outcomes

To teach understanding of algorithmic and mathematical principles behind numerical realizations of selected statistical methods, as well as behind certain numerical methods using statistical and probabilistic ideas. At the same time, cultivate skills in the implementations of those.

Assignment Project Exam Help

In particular, this is *neither*:

https://powcoder.com

- a programming course
    (you should have taken one in your 1st or 2nd year)

Add WeChat powcoder

- a specific programming language course
    (once you learn one, the second one goes quickly,
    and the third one even quicker)

- a general numerical mathematics course
    (there may be some overlap, but scope is different)

Questions? (But better let us move on to other choices)
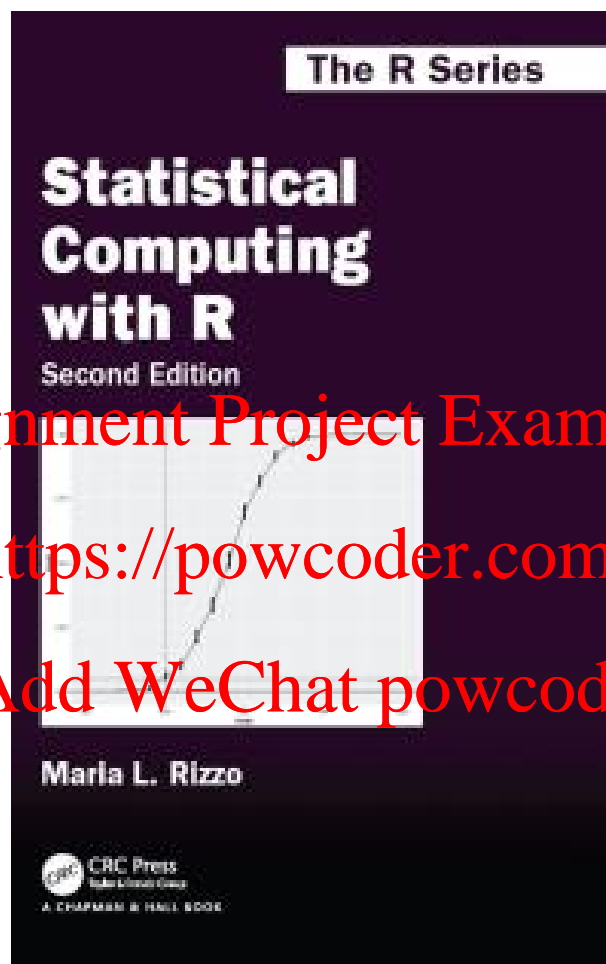
The textbook: NOT THIS ONE

Assignment Project Exam Help

https://powcoder.com

Add WeChat powcoder

(this is First Edition! Not terribly different, but still)

The textbook: but THIS ONE

The R Series

Statistical Computing with R

Second Edition

Assignment Project Exam Help

https://powcoder.com

Add WeChat powcoder

Maria L. Rizzo

CRC Press

A CHAPMAN & HALL BOOK

Maria L. Rizzo: Statistical Computing with R,
Second Edition, CRC Press 2019.
(Check out `www.library.ualberta.ca`)

Prerequisites: mathematics

Well, you need some…

1. Linear algebra. Every scientific computation starts with matrix algebra; every language facilitating this should be well-equipped for it.

Questions to ask yourself. Do you know what matrix is? Are you capable of multiplying them? Are you capable to move around in expressions involving vectors and matrices; for instance, do you know how transposition and inversion deal with multiplication? Do you know which matrices are symmetric, upper triangular, positive definite? What is inner product, how to write that neatly and how to express a vector norm through it? What is a determinant of a matrix? Have you ever handled operations with matrices divided to blocks?

(Rizzo is touching this only indirectly, if at all.)

Prerequisites: mathematics (continued)

2. (At least some) multivariate calculus really necessary…

Does a word "gradient" sound new to you? Can you manipulate partial derivatives and handle the related notation? Also the *second* partial derivatives? Exchangeability of the order of derivatives - familiar with? A word "Hessian"?

(The coverage will be quite a bit beyond the cookbook approach of Chapter 14 of Rizzo.)

3. Last but not least, some elementary, but important math will be used as well. For instance, does notation T^{-1} mean exclusively $1/T$ for you? Or also something else?

Prerequisites: probability

(Rizzo's book is being praised for providing some review of probability and statistics; however, it is a bit sketchy, and no exercises are available.)

It is *absolutely necessary* to know what the definitions in Section 2.1 mean and, in particular, *how to work with them*: rules on page 41 should be well known and digested, as they are to be actively used

Section 2.5, and quite a bit more beyond, will be covered in lectures. Regarding theorems like this, quite more important than memorizing the theorem it is important to *understand* their assumptions and outcomes. (We will try together.)

Section 2.8 gives a very brief review of Markov chains - those are needed to understand

Prerequisites: statistics

These topics will be reviewed when they become needed; Section 2.6 (maximum likelihood) and 2.7 (Bayesian analysis) may be helpful (although hardly self-contained) then

Again, most important here is to understand *the rules of game* here; that can take some time - and much more than just learning the theorems, proofs, and how to solve the exercises.

First homework will be posted soon...

Also - we are going to return to that shortly, but now:

Necessities

- Access to computer(s) (where you can install all the below)

  e.g. own laptop

- Internet access (preferably with Google within reach)

- A working installation of R

  possibly with a GUI (like that for OS X on Mac)

  or within an IDE (like ESS within Emacs or RStudio)

- A *programming* editor *of your choice*:

  (once an IDE, you have it automatically there)

  otherwise Vim? Atom? Notepad? TextEdit? Gedit?

  an important feature: with a formatting extension for code
  (Emacs has ESS, and RStudio is dedicated for R)

Finally, you will also need

- A *document* editor *of your choice*:

  e.g. TeX (LaTeX)

  or Microsoft Word? Google Drive?

Any questions?