Political Analysis Using R

Article *in* Journal of statistical software · October 2016

DOI: 10.18637/jss.v074.b01

1 author:

Some of the authors of this publication are also working on these related projects:

Project   Social Benefits and Costs of the National Flood Insurance Program View project

Project   Crystal Cube View project

Reviewer: James P. Howard, II
University of Maryland University College

Political Analysis Using R

James E. Monogan III's new book, *Political Analysis Using R*, is a brief text providing a strong outline for approaching political analysis using R. At the outset, it is important to understand this text describes political analysis in the sense of political science, public policy, and to a lesser degree, applied economics and macrosociology. This book does not provide methods or a package for performing political analysis from an electioneering standpoint. With that clarification out of the way, we can consider the text and its contents.

The first three chapters present the standard introduction to the R programming language. The first chapter presents a basic overview of how to download and install R, along with the various add-on packages available via CRAN. The second chapter goes further with examples of how to load data, read the basic attributes of a data set, along with data cleansing methods. The third chapter provides an introduction to visualizing data using R. Together, these chapters provide the fundamental introduction to R that a reader who may not be familiar with the language can use to get started. Further, owing to the political analysis aims, the examples given focus on political science data, such as human rights observations or the cost of oil.

Following the basics, the text moves into elementary statistical methods. The next chapter provides descriptive statistics and Chapter 5 provides methods for inference and bivariate analysis. This includes two-sample tests and cross-tabulations using the χ^2 test. Chapters 6 and 7 advance the statistical front to provide basic linear regression models, and generalized linear models. Political scientists will be happy to see that the logit and probit models are each given a full discussion, along with a combined interpretation discussion, focusing on the role of the link function. Briefer introductions to ordinal and count models are also provided. Despite the brevity of these reviews, they are sufficient within the generalized linear model framework presented.

The final third of the book pushes substantially further into statistical analysis, and a bit beyond. Chapter 8 introduces more methods for analysis using packages, such as multilevel regression and Bayesian regression. As Bayesian analysis continues to reshape political analysis, and many full-length treatments of Bayesian analysis in political science are available,

these six pages provide the briefest taste of the Bayesian methods. Chapter 9 pushes in a different direction, introducing time series methods. Like Bayesian analysis, this is well covered in many texts, and only a brief introduction to using the **pacf** and **TSA** packages is provided. While somewhat old fashioned, we might like to see the **xts** and **zoo** packages discussed here, where they are not.

The text ends with two final chapters focusing on the language R in greater detail. Chapter 10 gives us a chapter-length lesson on linear algebra in R. This includes a review of matrices and vectors, accessing the elements of vectors and matrices, and some basic matrix algebra. This is followed by a demonstration of calculating the ordinary least squares using the matrix methods. As a review of linear algebra, this is nice, but not necessarily advantageous to the reader who is interested in statistical methods. The final chapter provides a review of R's internal control structures. Finally, two sections discuss object-oriented programming in R and an elementary Monte Carlo simulation.

We should first discuss the drawbacks of this text. While detailed, they are not fatal to the text. The time series methods presented are out of step with the every day practice of analysts in the field. Further, it seems excessive to include linear algebra methods, without the greater context of numerical methods to support the discussion. Finally, the discussion of object-oriented programming is probably not necessary, either. The reason these are substantial drawbacks relates to the best audience for this text: a companion for introductory upper- or graduate-level courses in statistical methods for political analysis. For that audience, the linear algebra and object-oriented programming examples will not carry much weight.

Paired with a core text on statistical methods, either generalist or for applied political science or public policy, Monogan's book is an excellent addition to a political methodology classroom. The paired text would be able to provide background on the statistical methods and some motivation for how and why the methods work. Then, because nobody actually calculates statistical methods by hand, *Political Analysis Using R* can show the student how to apply the methods using R. As universities seek to push down the costs of textbooks, basing a statistical course in the free R drives down the cost over Stata or SAS. But without the core statistical discussion, Monogan's text is insufficient for a course.

The problems in Monogan's book are somewhat insufficient for application in a classroom. Each of the problems are themselves strong and well-structured for demonstrating learning. However, the number is simply insufficient to support a classroom, which tends to require a number of problems, either to split amongst the students or for multiple problems per student. The number, four to six per chapter, is just not enough for a classroom. This further pushes the need for a core statistical methodology text to pair Monogan's book with in a course. However, for an independent learner already well-versed in both statistical and political methodology, the problems can be enough to give the reader a testing ground for clarifying the examples shown.

Finally, we should mention the vast data available to support the text. Data for political analysis is readily available, and Monogan has taken full advantage of this. Examples using survey data, official records, and economic data support the analytical methods. In addition, both the examples in the text and the supporting data are available from Springer's website and the Dataverse website.

Overall, the text has its place, and evaluated within that place, the book is a strong candidate for classroom adoption. If I were teaching a course in political science or public policy

methodology, and using R to support the course, I would not hesitate to assign this book. The book provides an excellent overview of the application of statistical methods using R at a price that is less than the student-priced edition of Stata. That is a major benefit for both students and professors concerned about the ever rising cost of educational materials.

Reviewer:

James P. Howard, II
Mathematics and Statistics
University of Maryland University College
E-mail: jh@jameshoward.us or jphoward@faculty.umuc.edu
Twitter: @howardjp
URL: http://jameshoward.us