# Book Review: Spatial Statistics and Geostatistics

Blog Admin

This book aims to explain and demonstrate techniques in spatial sampling, local statistics, and advanced topics including Bayesian methods, Monte Carlo simulation, error and uncertainty. **Spatial Statistics and Geostatistics** *is highly recommended to researchers in geography, environmental science, health and epidemiology, population and demography, and planning, writes* **Hamish Clift**.

Spatial Statistics and Geostatistics. Yongwan Chun and Daniel A. Griffith. SAGE. January 2013.

Find this book:

It is a tired but relevant cliche that the world is getting smaller and smaller, driven by processes of globalisation and the evolution of information technology. As this occurs, it is necessary to place scientific data within a location-based frame of reference in order to make better informed decisions regarding public health, economics and political policy. By adding meaningful location data to a set of statistics, the dividend of research can be increased by orders of magnitude. In a new book, Yongwan Chun and Daniel A. Griffith, both scholars at the University of Texas at Dallas, describe the importance of spatial statistics and geostatistics and provide an excellent introduction to the area for economists, epidemiologists, geographers and statisticians who are already able to draw upon knowledge and ability in the field of modern statistical analysis.
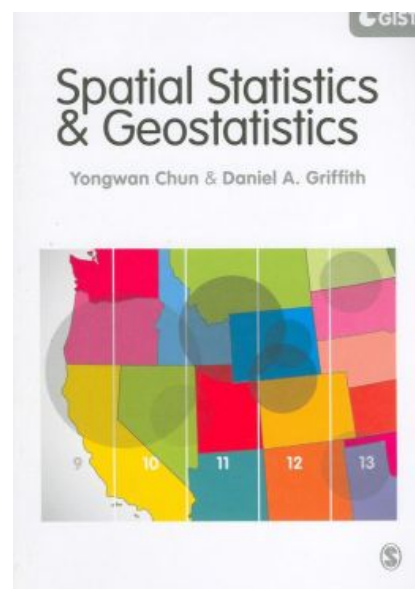
The book contains nine chapters that build upon one another, beginning with spatial autocorrelation, sampling and composition and configuration before moving on to chapters specifically related to econometrics, local statistics, spatial variance and covariance and spatial interpolation in two dimensions. It finishes with a chapter on advanced topics, which includes Bayesian methods, designing Monte Carlo simulation experiments and spatial error, although there are points throughout the book where the authors offer thoughts on specific topics to do with the practical applicability of this branch of statistics.

Each chapter also includes a section of code for the R project R is a command line driven suite of software for manipulating and analysing data which should be familiar to most readers of the book. Chun and Griffith chose to use R because it is freely distributed under a GNU license and has geographic information system capabilities as well as the capacity to run the techniques described in the book. The benefits to R are that it is modular and extensible and can be compiled and run on any platform, making it particularly accessible to students. The modularity of R means that it can be used for a variety of different purposes, which are enabled by simply downloading packages from the Comprehensive R Archive Network (CRAN) website.

The authors have collected an amazing gamut of data from the main island of Puerto Rico, including environmental, demographic, economic, census and geographical data. This data is used in the empirical examples at the end of each chapter, but the entire book relies upon the data to exemplify and illustrate the techniques it discusses to great effect. The continuity of the data throughout the book really helps to ground the reader and certainly increases understanding. Not to mention that it is actually really cool to be able to see first hand how these theories work in real (and relevant) situations.

After discussing the problem of spatial autocorrelation, Chun and Griffith move, in chapter 3, to deal with special sampling. In this chapter the authors describe techniques for collecting spatial data and implementing resampling techniques. They also take time to detail the concept of effective geographic sample size. The collection of Puerto Rican data is used to great effect in this chapter to illustrate the importance of designing a data collection system which is appropriate to the population and location of the data. The authors also demonstrate how to use random resampling techniques and describe the importance and necessity of using such techniques in various situations.

Chapter 6 shifts from considering global statistics – the broad theme of chapters 2 to 5 – to local geographic statistics and in doing so demonstrates the relationship between local and global. The chapter is concerned with local spatial relationships and the analysis of spatial dependency, the focus of which, the authors explain, is on differences rather than similarities. This, in turn, raises problems of homogeneity and autocorrelation.

The penultimate chapter of the book, "Methods for Spatial Interpolation in Two Dimensions", describes how statisticians can deal with sets of data which are incomplete due either the absence of observations in the dataset, or the absence of values for locations in a geographic landscape. There are lots of reasons for data to be less than ideal, particularly in fields such as infectious disease epidemiology and in development economics, where data gathering is fraught with problems. The authors proceed to explain a few ways of interpolating data where there are absences, including kriging (which is the preferable method) and the EM algorithm combined with spatial autocorrelation and Eigenvector spatial filtering.

Spatial Statistics and Geostatistics is an excellent resource for both the student and practitioner. It is clear and thoughtfully laid out, offers excellent analysis and insight into the different methods and problems that arise from performing spatial and geostatistical analysis. It is loaded with the real world examples made available by the Puerto Rican data to set and the use of the R project to run statistical analysis, coupled with the provided code makes the ideas and analyses easy to practice and quicker to grasp. There is an excellent academic rigour to the book, which is demonstrated by its continual references to seminal works which have informed and shaped the discipline from its genesis in the 1950s and evolution to the present day. This rigour further increases the value of the book to academics, students and practitioners. The benefits of working with location sensitive data are clear and there are great benefits to be gained from the understanding and application of spatial statistics and geostatistics to a wide variety of scientific disciplines. *Spatial Statistics and Geostatistics* is highly recommended.

_____

Hamish Clift is a student of Law at QUT in Australia. He holds a BA in literature and philosophy and a Master of Development Economics from the UQ. His research interests include law, economics, governance, development and ethics. He tweets, largely mundanely, at @perishablexcess. Read more reviews by Hamish.