



# Short Book Reviews

Editor: Simo Puntanen

## Statistical Methods for Disease Clustering

Toshiro Tango

Springer, 2010, x + 247 pages, €69.95/£59.99/\$79.95, hardcover

ISBN: 978-1-4419-1571-9

### Table of Contents

- |  |  |
|--|--|
| 1. Introduction  | 6. General tests for spatial clustering: case-control point data |
| 2. Clustering and clusters                                   | 7. Tests for space-time clustering                               |
| 3. Disease mapping: visualization of spatial clustering      | 8. Focused tests for spatial clustering                          |
| 4. Tests for temporal clustering                             | 9. Space-time scan statistics                                    |
| 5. General tests for spatial clustering: regional count data | A. List of R functions   |

*Readership:* Students and researchers in biostatistics, epidemiology and environmental sciences.

There is an increasing awareness nowadays of environmental health risks, including bio-terrorism, together with the development of modern data collection systems. This combination provides a major impetus for the sorts of studies described in this book. The basic question addressed is: are the observed events (of disease) compatible with random occurrence in space and time or is there some discernible pattern? We have all seen the headlines in recent years proclaiming the proximity to nuclear sites of outbreaks of childhood leukemia, or of various health issues in the vicinity of waste incinerators. However, it is often difficult to establish any link, partly because random events do sometimes cluster without any prompting. The book gives a survey of current statistical methods for detecting clustering, meaning, that, which is unlikely to be purely random, however the latter is defined. The methodology is almost all retrospective in the sense of examining past records of disease.

Chapter 1 is introductory, telling the reader what disease clustering is and how the book is organised. Particularly helpful here is Section 1.1, in which different types of disease clustering are defined and a passage is quoted from a relevant source illustrating a particular application for each. Chapters 2 and 3 cover the basic theory: spatial point processes, types of data, statistical models and inference. Chapters 4 to 9 conform to a fairly standard pattern. They begin with one or two examples, for which numerical data or geographical figures are presented. There follow the hypothesis to be tested, a historical review of statistical methods, a detailed description of selected methods, application to the data, and some further relevant discussion.

Overall, the book is very practically oriented. R-code is given throughout for implementing the various analyses, together with a website containing a set of R-functions, though not for downloading data sets as far as I can see. Prerequisites include some working knowledge of basic Statistical methods, particularly in Biostatistics and Epidemiology. Exercises are not included but one would be encouraged to try out the methods on any data to hand. The explanations of the methods are very clear and detailed and the discussion is informative. So, I believe that the book will be of good use to the novice coming into the subject and also for the initiated wishing to broaden their perspective.

Martin Crowder: [m.crowder@imperial.ac.uk](mailto:m.crowder@imperial.ac.uk)  
Mathematics Department, Imperial College  
London SW7 2AZ, UK