
Preface

The general perception of statistics is that of an extremely challenging field of endeavour, and if you do something that difficult, you must be a good scientist. In our courses we often encounter people saying, '*I have been to the statistics course given by Professor So and So, and it was very good. Lots of difficult subject matter...but I didn't understand any of it.*' That does not make much sense. In our opinion doing simple things is already sufficiently demanding.

In 2010 we wrote a paper on data exploration. The paper presented a simple protocol containing elementary common sense steps to be carried out before applying a statistical model. Hardly worth the time it took to write or to read, or so you would expect. This paper received the highest number of downloads in the history of the journal in which it was published and was identified by Sciencewatch.com from Thomas Reuters (the people who calculate journal impact factors) as a highly cited paper in ecology and biology. This shows that simplicity gets the attention.

So, forget about embarking upon gruelling statistical expeditions, do simple things! Of course, in one person's view, a given analytical procedure may be simple, while others may find it hopelessly complex. Many statistical methods are straightforward, once you understand the basic principle. The challenge is in gaining the necessary knowledge without digging through a mountain of mathematical equations and theorems.

The subject of this book is generalized additive modelling (GAM), which is perceived by some as difficult. Yet, it is essentially generalized linear modelling. We discussed GAM in our 2007, 2009^a, and 2012 books. Intermediate level explanations are given in the 2007 and 2009^a volumes, and in the 2012 book GAM is presented in a zero-inflation context.

This book provides an introduction at the basic level. The danger in writing a *Beginner's Guide* book is that it may become a sort of cookbook that distances the reader from any critical thinking. That is not the intent of this book. If you are able to program a statistical analysis, or at least its underlying principles, from scratch, you will grasp the rationale behind it. The nitty-gritty details of the implementation of a statistical method are something that you can take for granted; if you can perform the basics, you can understand it. The aim of this book is to take you through the world of GAM and show that it is a simple and useful statistical modelling technique.

This book is the first in a series of *Beginner's Guide* books that will be published by Highland Statistics staff and colleagues in 2013 and 2014:

- *A Beginner's Guide to Generalized Additive Mixed Models with R*. Zuur AF, Saveliev AA, Ieno EN
- *A Beginner's Guide to Generalized Linear Models with R and JAGS, OpenBUGS and WinBUGS*. Zuur AF, Hilbe JM, Ieno EN
- *A Beginner's Guide to Data Exploration and Visualisation with R*. Ieno EN, Zuur AF

Acknowledgements

The material contained in this book has been presented in a large number of courses between 2003 and 2012, and I am greatly indebted to all participants who helped improve the material. I would also like to thank those who read and commented on parts of earlier drafts: Mara Schmiing, Anatoly Saveliev, Elena Ieno, and Philip Dixon. I thank them all for their positive, encouraging, and useful reviews. Their comments and criticisms greatly improved the book. Special thanks to the contributors to this book who kindly donated data sets for the case studies and helped to make it possible to explain GAM. I would also like to thank Monty Priede for allowing use of the two data sets used in Chapters 2 - 4.

I also thank Kathleen Hills of The Lucidus Consultancy for editing this book.

A big ‘thank you’ is also due to the large number of folks who wrote R (R Development Core Team, 2012) and its many packages. I made extensive use of the `lattice` (Sarkar 2008) and `mgcv` (Wood 2006) packages. This thank you is also on behalf of the readers of this book.

Datasets used in this book

This book uses primarily biological data. Nevertheless, whatever your area of study and whatever your data, the procedures presented will apply. All datasets used in this book are downloadable from www.highstat.com.

Cover art

The cover drawing of oystercatchers was made by Jon Thompson (www.yellowbirdgallery.org). Mr Thompson was born in 1939 to Irish parents and has lived most of his life in Scotland. In the 1980s, he was instinctively drawn to the Orkney Islands and is continually inspired by the landscape and bird life of Orkney. He has been creating bird art for 30 years through drawing, painting, sculpture, and jewellery, never attempting to reproduce nature, but to draw parallels with it. A close up view of a bird feather is all the inspiration he needs.

Newburgh, UK & Mount Irvine, Tobago
Alain F. Zuur
September 2012