

## Preface

We have written this book for professional researchers outside the field of mathematics, people who do not spend their time wondering about the intricacies of generalizing a result from discrete space to  $\mathfrak{R}_1$  but who, nonetheless, understand statistics. Our readers may sometimes be sloppy when they say that a probability density is a probability, but when pressed, they know there is a difference and remember that a probability density can indeed even be greater than one. However, our readers are never sloppy when it comes to their science. Our readers use statistics as a tool, just as they use mathematics, and just as they sometimes use computer software.

This is a book about survival analysis for the professional data analyst, whether a health scientist, an economist, a political scientist, or any of a wide range of scientists who have found that survival analysis is applicable to their problems. This is a book for researchers who want to understand what they are doing and to understand the underpinnings and assumptions of the tools they use; in other words, this is a book for all researchers.

This book grew out of software, but nonetheless, it is not a manual. That genesis, however, gives this book an applied outlook that is sometimes missing from other works. We, the authors of this book, are also the authors of Stata's survival analysis commands, which have had something more than modest success. Writing application software places a discipline on authors not unlike that of building of scientific machines by engineers. Problems that might be swept under the rug as mere details cannot be ignored, in the construction of software, and the authors are often reminded that the devil is in the details. It is those details that cause users such grief, such confusion, and sometimes, such pleasure.

In addition to having written the software, we have all been involved in supporting it, which is to say, interacting with users (real professionals). We have seen the software used in ways that we would never have imagined, and we have seen the problems that arise in such uses. Those problems are often not simply programming issues but involve statistical issues that have given us pause. To the statisticians in the audience, we mention that there is nothing like embedding yourself in the problems of real researchers to teach you that problems you thought unimportant are of great importance, and vice versa. There is nothing like "straightforwardly generalizing" some procedure to teach you that there are subtle issues worth lots of thought.

In this book, we illustrate the concepts of using Stata. Readers should expect a certain bias on our part, but the concepts go beyond our implementation of them. We

will often discuss substantive issues right in the midst of issues of computer use, and we do that because, in real life, that is where they arise.

This book also grew out of a course we taught several times over the web, and the many researchers who took that course will find in this book the companion text they lamented not having for that course.

We do not wish to promise more than we can deliver, but the reader of this book should come away not just with an understanding of the formulas, but an intuition of how the various survival analysis estimators work and exactly what information they exploit.

We would like to thank all the people who over the years have contributed to our understanding of survival analysis and the improvement of Stata's survival capabilities be it through programs, comments, or suggestions. We are particularly grateful to

David Clayton of the Cambridge Institute for Medical Research  
Joanne M. Garrett of the University of North Carolina  
Michael Hills retired from the London School of Hygiene and Tropical Medicine  
David Hosmer, Jr. of the University of Massachusetts, Amherst  
Stephen P. Jenkins of the University of Essex  
Stanley Lemeshow of Ohio State University  
Adrian Mander of the MRC Biostatistics Unit  
William H. Rogers of The Health Institute at New England Medical Center  
Patrick Royston of the MRC Clinical Trials Unit  
Peter Sasieni of Cancer Research UK  
Jeroen Weesie of Utrecht University

By no means is this list complete; we would like to express our thanks as well to all those who should have been listed.

Mario A. Cleves  
William W. Gould  
Roberto G. Gutierrez

College Station, Texas  
May 2002