# Preface

Survival and event history analysis have developed into one of the major areas of biostatistics, with important applications in other fields as well, including reliability theory, actuarial science, demography, epidemiology, sociology, and econometrics. This has resulted in a number of substantial textbooks in the field. However, rapidly developing theoretical models, combined with an ever-increasing amount of high-quality data with complex structures, have left a gap in the literature. It has been our wish to provide a book that exposes the rich interplay between theory and applications. Without being unnecessarily technical, we have wanted to show how theoretical aspects of statistical models have direct, intuitive implications for applied work. And conversely, how apparently disparate and complex features of data can be set in a theoretical framework that enables comprehensive modeling.

Most textbooks in survival analysis focus on the occurrence of single events. In actual fact, much event history data consist of occurrences that are repeated over time or related among individuals. The reason for this is, in part, the development toward increased registration and monitoring of individual life histories in clinical medicine, epidemiology, and other fields. The material for research and analysis is therefore more extensive and complex than it used to be, and standard approaches in survival analysis are insufficient to handle this.

It is natural to view such life histories as stochastic processes, and this is the basic idea behind our book. We start with the now classical counting process theory and give a detailed introduction to the topic. Leaving out much mathematical detail, we focus on understanding the important ideas. Next, we introduce standard survival analysis methods, including Kaplan-Meier and Nelson-Aalen plots and Cox regression. We also give a careful discussion of the additive hazards model. Then we extend the use of the counting process framework by counting several events for each individual. This yields very fruitful models, especially when combined with the additive hazards model. We further include time-dependent covariates, or marker processes, to define what we term dynamic path analysis, an extension of classical path analysis to include time. This allows us to explicitly analyze how various processes influence one another. Most of this is new, or very recently developed, material.

Another new aspect of the book is the explicit connection drawn between event history analysis and statistical causality. The focus is on causal formulations where time is explicitly present, including ideas like Granger causality, local independence, and dynamic path analysis.

Unique to this book is the emphasis on models that give insight into the rather elusive concept of hazard rate, and the various shapes the hazard rate can have. The effect of unobserved heterogeneity, or frailty, is broadly discussed with focus on a number of artifacts implied by the frailty structure as well as on applying these models to multivariate survival data. A new class of process-based frailty models is introduced, derived from processes with independent increments. We also study models of underlying processes that, when hitting a barrier, lead to the event in question. It is emphasized that frailty is an essential concept in event history analysis, and wrong conclusions may be drawn if frailty is ignored.

The applied aspect of the book is supported by a range of real-life data sets. Most of the data are well known to us through active collaboration with research groups in other fields. By necessity, some of the examples based on them have been simplified to achieve a more pedagogical exposition. However, we have made a strong effort to keep close to the relevant research questions; in particular, our three case studies are intended as more or less full-blown analyses of data of current interest.

Hence, the book contains a lot of new material for both the practicing statistician and those who are more interested in the theory of the subject. The book is intended primarily for researchers working in biostatistics, but it should also be found interesting by statisticians in other fields where event history analysis is of importance. The reader should have some theoretical background in statistics, although the mathematics is kept at a "user" level, not going into the finer theoretical details.

The book should be well suited as a textbook for graduate courses in survival and event history analysis, in particular for courses on the analysis of multivariate or complex event history data. Some exercises are provided at the end of each chapter, and supplementary exercises, as well as some of the datasets used as examples, may be found on the book's Web page at www.springer.com/978-0-387-20287-7.

A number of friends and colleagues have read and commented on parts of this book, have participated in discussion on themes from the book, or have provided us with data used as examples, and we want to thank them all: Per Kragh Andersen, Vanessa Didelez, Ludwig Fahrmeir, Johan Fosen, Axel Gandy, Jon Michael Gran, Nina Gunnes, Robin Henderson, Nils Lid Hjort, Niels Keiding, Bryan Langholz, Stein Atle Lie, Bo Lindqvist, Torben Martinussen, Sven Ove Samuelsen, Thomas Scheike, Finn Skårderud, Anders Skrondal, Halvor Sommerfelt, Hans Steinsland, Hans van Houwelingen, and Stein Emil Vollset.

Hege Marie Bøvelstad and Marion Haugen did an invaluable job making graphs for a number of the examples. They also fearlessly took on the daunting task of preparing all our references, saving us long hours of tedious work.

Preliminary versions of the book have been used as the text for a graduate course in survival and event history analysis at the University of Oslo and for a Nordic course organized as part of a Nordplus program for master-level courses in

biostatistics. We thank the students at these courses, who gave valuable feedback and pointed out a number of typos.

Most of the work on the book has been done as part of our regular appointments at the University of Oslo (Odd O. Aalen and Ørnulf Borgan) and the Norwegian Institute of Public Health (Håkon K. Gjessing), and we thank our employers for giving us time to work on the book as part of our regular duties. In addition, important parts of the work were done while the authors were participating in an international research group at the Centre for Advanced Study at the Norwegian Academy of Science and Letters in Oslo during the academic year 2005–2006. We express our gratitude to the staff of the center for providing such good working facilities and a pleasant social environment.

Oslo                                                                                      *Odd O. Aalen*
January 2008                                                                        *Ørnulf Borgan*
                                                                                  *Håkon K. Gjessing*