Statistical Models and Methods for Lifetime Data

Second Edition

JERALD F. LAWLESS
University of Waterloo

A JOHN WILEY & SONS, INC., PUBLICATION

This page intentionally left blank

Statistical Models and Methods for Lifetime Data

WILEY SERIES IN PROBABILITY AND STATISTICS

Established by WALTER A. SHEWHART and SAMUEL S. WILKS

Editors: *David J. Balding, Peter Bloomfield, Noel A. C. Cressie, Nicholas I. Fisher, Iain M. Johnstone, J. B. Kadane, Louise M. Ryan, David W. Scott, Adrian F. M. Smith, Jozef L. Teugels*
Editors Emeriti: *Vic Barnett, J. Stuart Hunter, David G. Kendall*

A complete list of the titles in this series appears at the end of this volume.

Statistical Models and Methods for Lifetime Data

Second Edition

JERALD F. LAWLESS
University of Waterloo

A JOHN WILEY & SONS, INC., PUBLICATION

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data

Lawless, J. F. (Jerald F.), 1944–
Statistical models and methods for lifetime data / Jerald F. Lawless.– 2nd ed.
p. cm. – (Wiley series in probability and statistics)
Includes bibliographical references and index.
ISBN 0-471-37215-3 (cloth : acid-free paper)
1. Failure time data analysis. 2. Survival analysis (Biometry). I. Title. II. Series.
QA276 .L328 2003
519.5—dc21 2002151805

Printed in the United States of America

10 9 8 7 6 5 4 3 2

To Liz

This page intentionally left blank

Contents

Preface to the Second Edition

The modeling and analysis of lifetimes is an important aspect of statistical work in a wide variety of scientific and technological fields. This book provides a comprehensive account of models and methods for lifetime data.

The field of lifetime data analysis has grown and matured since the first edition of the book was published. This second edition has accordingly been rewritten to reflect new developments in methodology, theory, and computing. The orientation and philosophy, however, remain close to the first edition. Lifetime data analysis is covered without concentrating exclusively on any specific field of application, though as before most of the examples are drawn from engineering and the biomedical sciences. There is a strong emphasis on parametric models, but non- and semiparametric methods are also given detailed treatments. Likelihood-based inference procedures are emphasized and serve to unify the methodology; implementation using both special lifetime data software, and general optimization software is discussed.

Extensive developments in software have made it possible to focus less on computational details and simplified methods of estimation than in the first edition, and to expand examples and illustrations. Graphical tools now feature more prominently. Many new references have been added, and some references and material from the first edition thus have been dropped. I have attempted, however, to retain enough early references to indicate the origins and evolution of different topics.

This edition of the book, like the first, is intended to serve as a reference for individuals interested in the analysis of lifetime data and as a text for graduate courses. Several appendices review aspects of statistical theory and computation that underlie the methodology presented in the book. A Problems and Supplements section concludes each chapter. There are many statistical software packages with capabilities in lifetime data analysis, and I have not attempted to provide an overview or comparison. Most of the examples were prepared using S-Plus, but other packages could have been used. Brief Computational Notes are provided at the ends of most chapters. Data sets discussed in the book are almost all either given or identified as being available electronically from Web locations mentioned in Appendix G.

There has been a small reorganization of topics from the first edition, consisting mainly of an expanded discussion of observation schemes and censoring (now in a new Chapter 2), and an expanded treatment of multiple failure modes (now given in

a separate Chapter 9). Several new topics have been introduced, including counting process-martingale tools, resampling and simulation methodology, estimating function methods, treatments of interval censored and truncated data, and discussions of multivariate lifetimes and event history models. In addition, material on many other topics has been updated and extensively revised, as have the Problems and Supplements sections.

To keep the book at a reasonable length I have had to omit or merely outline certain topics that might have been included. Some such topics are mentioned in the Bibliographic Notes that have been introduced at the ends of chapters, or in the Problems and Supplements sections. Statistical science encyclopedias (e.g. Kotz et al. 1988; Armitage and Colton 1998) are valuable sources of further information on a wide range of topics associated with lifetime data. Web-based tools for locating resource materials also make it relatively easy to research topics not covered in the book.

Chapter 1 contains introductory material on lifetime distributions and surveys important models. Chapter 2 deals with observation schemes for lifetime data and the formation of likelihood functions. Chapter 3 discusses graphical methods and nonparametric estimation of distribution characteristics based on different types of lifetime data. Chapter 4 introduces inference procedures for parametric models, including exponential, gamma, inverse Gaussian, and mixture models. Chapter 5 provides corresponding procedures for log-location-scale models and extensions to them; the Weibull, log-normal and log-logistic models are treated in detail. Chapter 6 discusses regression models, exploratory and diagnostic methods, and develops inference procedures for parametric models. Chapter 7 deals with semiparametric methodology for proportional or multiplicative hazards models. Chapter 8 presents rank-based and semiparametric procedures based on location-scale models. Chapter 9 gives a thorough treatment of multiple failure modes, or competing risks. Chapter 10 discusses goodness-of-fit tests and describes procedures for specific models in the book. Finally, Chapter 11 introduces several important topics that go beyond univariate survival analysis: multivariate lifetime models, sequences of lifetimes, event history processes, and joint models for lifetimes and coprocesses. It is shown how the methods of previous chapters can be applied to many problems in these areas.

I am indebted to various individuals for their contributions to this edition of the book. Ker-Ai Lee and Melanie Wigg assisted with computing and the preparation of examples. Some examples are based on joint work with Richard Cook, Jack Kalbfleisch, and graduate student Wenqing He. I have benefitted for many years from collaboration and conversations with Richard Cook, Jack Kalbfleisch, and Jock MacKay, and from my interactions with numerous fine graduate students at Waterloo.

I want to acknowledge and thank Lynda Clarke, who has labored long, hard, and expertly on the manuscript, as she did on the first edition of the book 20 years ago.

The University of Waterloo's Department of Statistics and Actuarial Science has provided a stimulating environment for research and teaching throughout my career. Part of the work for this edition was done during a sabbatical leave spent at University

of Auckland (January–March 2000) and at University College London (January–March 2001); their hospitality is gratefully acknowledged. I also want to acknowledge support over many years from the research grants programs of the Natural Sciences and Engineering Research Council of Canada (NSERC), and to thank General Motors Canada for their cosponsorship, with NSERC, of a personal Industrial Research Chair.

Finally, I thank my family and especially my wife, Liz, for her patience and support during this project.

JERALD. F. LAWLESS

Waterloo, Ontario
April 2002

This page intentionally left blank

Preface to the First Edition

The statistical analysis of lifetime or response time data has become a topic of considerable interest to statisticians and workers in areas such as engineering, medicine, and the biological sciences. The field has expanded rapidly in recent years, and publications on the subject can be found in the literatures of several disciplines besides statistics. This book draws together material on the analysis of lifetime data and gives a comprehensive presentation of important models and methods.

My aim is to give a broad coverage of the area without unduly concentrating on any single field of application. Most of the examples in the book, however, come from engineering or the biomedical sciences, where these methods are widely used. The book contains what I feel are the most important topics in lifetime data methodology. These include various parametric models and their associated statistical methods, nonparametric and distribution-free methods, and graphical procedures. To keep the book at a reasonable length I have had to either sketch or entirely omit topics that could have usefully been treated in detail. Some of these topics are referenced or touched upon in the Problems and Supplements sections at the ends of chapters.

This book is intended as a reference for individuals interested in the analysis of lifetime data and can also be used as a text for graduate courses in this area. A basic knowledge of probability and statistical inference is assumed, but I have attempted to carefully lay out the models and assumptions upon which procedures are based and to show how the procedures are developed. In addition, several appendices review statistical theory that may be unfamiliar to some readers. Numerical illustrations are given for most procedures, and the book contains numerous examples involving real data. Each chapter concludes with a Problems and Supplements section, which provides exercises on the chapter material, and supplements and extends the topics discussed. For the reader interested in research on lifetime data methodology I have given fairly extensive references to recent work and outstanding problems.

Chapter 1 contains introductory material on lifetime distributions and surveys the most important parametric models. Censoring is introduced, and its ramifications for statistical inference are considered. In Chapter 2 some methods of examining univariate lifetime data and obtaining nonparametric estimates of distribution characteristics are discussed; life tables and graphical procedures play key roles. Chapters

3, 4, and 5 deal with inference for important parametric models, including the exponential, Weibull, gamma, log-normal, and generalized gamma distributions. This is extended in Chapter 6 to problems with concomitant variables, through regression models based on these distributions. Chapters 7 and 8 present nonparametric and distribution-free procedures: Chapter 7 deals with methods based on the proportional hazards regression model, and Chapter 8 gives distribution-free procedures for single- and many-sample problems. Goodness-of-fit tests for lifetime distribution models are considered in Chapter 9. Chapter 10 contains brief discussions of two important topics for which it was not feasible to give extended treatments: multivariate and stochastic process models. Several sections in this book are marked with asterisks; these contain discussions of a technical nature and can be omitted on a first reading.

A final remark concerning the methods presented is that the computer is, as always in modern statistics, a useful if not indispensible tool. For some problems, methods that do not require a computer are available, but more often access to a computer is a necessity. I have commented, wherever possible, on the computational aspects of procedures and have included additional material on computation in the Appendices.

Part of the work for the book was done during a sabbatical leave spent at Imperial College, London, and the University of Reading from 1978 to 1979; their hospitality is gratefully acknowledged. I would also like to express my appreciation to two extremely fine typists, Annemarie Nittel and Lynda Hohner, who labored long and diligently in the preparation of the manuscript.

J. F. LAWLESS

Waterloo, Ontario
June 1981

Statistical Models and Methods for Lifetime Data

This page intentionally left blank

CHAPTER 1

Basic Concepts and Models

1.1 INTRODUCTION

The statistical analysis of what are variously referred to as lifetime, survival time, or failure time data is an important topic in many areas, including the biomedical, engineering, and social sciences. Applications of lifetime distribution methodology range from investigations of the durability of manufactured items to studies of human diseases and their treatment. Some methods of dealing with lifetime data are quite old, but starting about 1970 the field expanded rapidly with respect to methodology, theory, and fields of application. Software packages for lifetime data analysis have been widely available since about 1980, with the frequent appearance of new features and packages.

This book presents and illustrates statistical methods for modeling and analyzing lifetime data. The aim is to provide a general treatment, and not focus exclusively on a particular field of application. Lifetime distribution methodology is widely used in the biomedical and engineering sciences, however, and most of the examples in the book come from those areas.

Throughout the book various types of data will, for convenience, be referred to as "lifetime" data. Basically, however, we consider situations in which the time to the occurrence of some event is of interest for individuals in some population. Sometimes the events are actual deaths of individuals and "lifetime" is the length of life measured from some particular starting point. In other instances "lifetime" and the words "death" or "failure," which denote the event of interest, are used in a figurative sense. In discussing applications, other terms such as "survival time" and "failure time" are also frequently used.

The following examples illustrate some ways in which lifetime data arise.

Example 1.1.1. Manufactured items with mechanical or electronic components are often subjected to life tests in order to obtain information on their durability. This involves putting items in operation, often in a laboratory setting, and observing them until they fail. It is common here to refer to the lifetimes as "failure times," since when an item ceases operating satisfactorily, it is said to have "failed."

Example 1.1.2. Demographers and social scientists are interested in the duration of certain life "states" for humans. Consider, for example, marriage and, in particular, the marriages formed during the year 1980 in a particular country. Then the lifetime of a marriage would be its duration; a marriage may end due to annulment, divorce, or death.

Example 1.1.3. In medical studies dealing with potentially fatal diseases one is interested in the survival time of individuals with the disease, measured from the date of diagnosis or some other starting point. For example, it is common to compare treatments for a disease at least partly in terms of the survival time distributions for patients receiving the different treatments.

Example 1.1.4. A standard experiment in the investigation of carcinogenic substances is one in which laboratory animals are subjected to doses of the substance and then observed to see if they develop tumors. A main variable of interest is the time to appearance of a tumor, measured from when the dose is administered.

The definition of lifetime includes a time scale and time origin, as well as a specification of the event (e.g., failure or death) that determines lifetime. In some settings it is difficult to say precisely when the event occurs: for example, this is the case for the appearance of a tumor in Example 1.1.4. The time scale is not always real or chronological time, especially where machines or equipment are concerned. For example, miles driven might be used as a time scale with motor vehicles, and number of pages of output for a computer printer or photocopier.

The main problems addressed in this book are those of specifying models to represent lifetime distributions and of making inferences based on these models. The objectives of modeling and statistical analysis include description or estimation of distributions, comparison of distributions, furthering scientific understanding, process or system improvement, prediction, and decision. Covariates or explanatory variables that can be related to lifetime usually feature prominently in these activities. In some settings there may be more than one lifetime variable associated with an individual, or an individual may die in different ways. The types of models used in lifetime data analysis range from fully parametric to nonparametric; semiparametric models that have both parametric and nonparametric features are common. The remaining sections of this chapter introduce lifetime models, but first we discuss some additional features and examples of lifetime data.

The chronological time needed to observe the lifetimes of all individuals in a study may be large enough that practical constraints prevent full observation. This leads to what is termed "censoring," in which an individual's lifetime is known only to exceed a certain value. In Example 1.1.1, for example, a life test might be terminated after, say, 28 days; if an item had not failed by that time, we would know only that its lifetime exceeded 28 days and refer to that value as a "censoring time." More generally, it may not be possible to determine exactly when a failure or death occurs, because individuals are seen only at certain times. In that case, we may know only that a lifetime lies in some interval (L, R); we refer to this as "interval censoring."

Another complication is that covariates associated with lifetime data may vary over time, and it may not be possible to observe their values at all times.

These and other features associated with lifetime data create interesting problems for analysis, and much of the development of the subject has been devoted to dealing with them. Chapter 2 considers these issues in detail. The remainder of this chapter covers the basic concepts of lifetime distributions and introduces important models. Section 1.2 discusses lifetime distributions generally, and Section 1.3 introduces important parametric univariate models. Sections 1.4 to 1.6 discuss more complex models involving covariates, multiple lifetimes, and multiple types of failure. Before turning to this, we consider a few examples of lifetime data, to illustrate some of the points just mentioned.

Example 1.1.5. Nelson (1972a) described the results of a life test experiment in which specimens of a type of electrical insulating fluid were subjected to a constant voltage stress. The length of time until each specimen failed, or "broke down," was observed. Table 1.1 gives results for seven groups of specimens, tested at voltages ranging from 26 to 38 kilovolts (kV).

The main purpose of the experiment was to investigate the distribution of time to breakdown for the insulating fluid and to relate this to the voltage level. Quite clearly, breakdown times tend to decrease as the voltage increases. In addition to the formulation of a model relating breakdown times and voltage, the estimation of the breakdown time distribution at a "normal" voltage of 20 kV was important. Breakdown times tend to be very large at 20 kV, and this involves a substantial extrapolation from the experimental data.

The experiment in Example 1.1.5 was run long enough to observe the failure of all the insulation specimens tested. Sometimes it may take a long time for all items to fail, and it is deemed necessary to terminate a study before this can happen. In this case, the lifetimes of certain items are censored. For example, if a decision had been

Table 1.1. Times to Breakdown (in minutes) at Each of Seven Voltage Levels

Voltage Level (kV)	n_i	Breakdown Times
26	3	5.79, 1579.52, 2323.7
28	5	68.85, 426.07, 110.29, 108.29, 1067.6
30	11	17.05, 22.66, 21.02, 175.88, 139.07, 144.12, 20.46, 43.40, 194.90, 47.30, 7.74
32	15	0.40, 82.85, 9.88, 89.29, 215.10, 2.75, 0.79, 15.93, 3.91. 0.27, 0.69, 100.58, 27.80, 13.95, 53.24
34	19	0.96, 4.15, 0.19, 0.78, 8.01, 31.75, 7.35, 6.50, 8.27, 33.91, 32.52, 3.16, 4.85, 2.78, 4.67, 1.31, 12.06, 36.71, 72.89
36	15	1.97, 0.59, 2.58, 1.69, 2.71, 25.50, 0.35, 0.99, 3.99, 3.67, 2.07, 0.96, 5.35, 2.90, 13.77
38	8	0.47, 0.73, 1.40, 0.74, 0.39, 1.13, 0.09, 2.38

Table 1.2. Lifetimes for 10 Pieces of Equipment

Item Number	1	2	3	4	5	6	7	8	9	10
Date of installation	11 June	21 June	22 June	2 July	21 July	31 July	31 July	1 Aug	2 Aug	10 Aug
Date of failure	13 June	—	12 Aug	—	23 Aug	27 Aug	14 Aug	25 Aug	6 Aug	—
Lifetime (days)	2	≥ 72	51	≥ 60	33	27	14	24	4	≥ 21

made in the preceding experiment to terminate testing after 180 minutes had elapsed, then two of the observations in the 26- and 28-kV sample and one each in the 30- and 32-kV samples would have been censored. In each case, we would not know the exact failure time of the item, but only that it exceeded 180 minutes.

Censoring arises in lifetime data in a variety of ways and is discussed in detail in Chapter 2. The remaining examples in this section all involve censoring of some kind.

Example 1.1.6. Bartholomew (1957) considered a situation in which pieces of equipment were installed in a system at different times. At a later date some of the pieces had failed and the rest were still in use. With the aim of studying the lifetime distribution of the equipment, Bartholomew gave the data in Table 1.2 for 10 pieces of equipment. The first item was installed on June 11 and data were collected up to August 31. At that time, three items (numbers 2, 4, and 10) had still not failed, and their failure times are therefore censored; we know for these items only that their failure times exceed 72, 60, and 21 days, respectively.

Example 1.1.7. Gehan (1965) and others have discussed the results of a clinical trial reported by Freireich et al. (1963), in which the drug 6-mercaptopurine (6-MP) was compared to a placebo with respect to the ability to maintain remission in acute leukemia patients. Table 1.3 gives remission times for two groups of 21 patients each, one group given the placebo and the other the drug 6-MP.

The starred observations are censoring times; for these patients, the disease was still in a state of remission at the end of the study. Censoring is common in clinical trials, since the trial is often terminated before all individuals have "failed." In addition, individuals may enter a study at various times, and hence may be under observation for different lengths of time. In this trial, individuals entered the study in matched pairs at different times and a sequential stopping rule was used to terminate the study (Klein and Moeschberger 1997, p. 2).

Example 1.1.8. Therneau and Hamilton (1997) discussed data that arose in a study of persons with cystic fibrosis (Fuchs et al. 1994). These individuals are susceptible to an accumulation of mucus in the lungs, which leads to pulmonary exacerbations and deterioration of lung function. In a clinical trial to investigate the efficacy of daily administration of a recombinant form of the human enzyme DNase 1 in preventing exacerbations, subjects were randomly assigned to the new treatment (called rhDNase or Pulmozyme) or a placebo. Subjects, who were exacerbation-free at ran-

Table 1.3. Lengths of Remission (in weeks) for Two Groups of Patients[a]

6-MP	6, 6, 6, 6*, 7, 9*, 10, 10*, 11*, 13, 16, 17*, 19*, 20*, 22, 23, 25*, 32*, 32*, 34*, 35*
Placebo	1, 1, 2, 2, 3, 4, 4, 5, 5, 8, 8, 8, 8, 11, 11, 12, 12, 15, 17, 22, 23

[a]Stars denote censored observations.

Table 1.4. Times to First Pulmonary Exacerbation for 10 Subjects

t (days)[a]	trt	fev[b]
168*	1	28.8
169*	1	64.0
65	0	67.2
168*	1	57.6
171*	0	57.6
166*	1	25.6
168*	0	86.4
90	0	32.0
169*	1	86.4
8	0	28.8

[a] Starred values are censoring times.
[b] fev measure is percent of predicted normal fev, based on sex, age, and height.

domization, were followed for approximately 169 days and the days at which an exacerbation period started were noted. When an exacerbation spell began, a subject was given antibiotics, and after the exacerbation had disappeared the subject was then considered at risk for a new exacerbation. Consequently, some subjects had no exacerbations over the 169-day follow-up period, some had one, and some had two or more.

There were 324 subjects assigned to the Placebo group by randomization, and 321 to the rhDNase group. The objective was to compare the two groups in terms of the avoidance of exacerbations. The simplest comparison is to note that 139 (43%) of Placebo subjects had at least one exacerbation and that 104 (32%) of rhDNase subjects did. A more comprehensive comparison can be based on the time t to the first exacerbation after randomization. Table 1.4 shows data for ten subjects: failure time t and two covariates, trt (= 0 for Placebo and 1 for rhDNase) and fev (forced expiratory volume at the time of randomization, which is a measure of initial pulmonary function). A still more comprehensive analysis might also use the data on second and subsequent exacerbations; this topic is discussed in Chapter 11.

Example 1.1.9. Table 1.5 presents survival data on 40 advanced lung cancer patients, taken from a study discussed by Prentice (1973). The main purpose of the study was to compare the effects of two chemotherapy treatments in prolonging survival time. All patients represented here received prior therapy and were then randomly assigned to one of the two treatments, termed "standard" and "test." Survival times t, measured from the start of treatment for each patient, are recorded in Table 1.5. Censored observations correspond to patients who were still alive at the time the data were collected. Concomitant variables that were thought possibly to be important are also shown for each patient. First, patients can have different types of tumors; they have been classified into four categories (squamous, small, adeno,