Statistics for Biology and Health

Borchers/Buckland/Zucchini: Estimating Animal Abundance: Closed Populations.
Everitt/Rabe-Hesketh: Analyzing Medical Data Using S-PLUS.
Ewens/Grant: Statistical Methods in Bioinformatics: An Introduction.
Hougaard: Analysis of Multivariate Survival Data.
Klein/Moeschberger: Survival Analysis: Techniques for Censored and Truncated Data, 2nd ed.
Kleinbaum: Survival Analysis: A Self-Learning Text.
Kleinbaum/Klein: Logistic Regression: A Self-Learning Text, 2nd ed.
Lange: Mathematical and Statistical Methods for Genetic Analysis, 2nd ed.
Manton/Singer/Suzman: Forecasting the Health of Elderly Populations.
Moyé: Multiple Analyses in Clinical Trials: Fundamentals for Investigators.
Parmigiani/Garrett/Irizarry/Zeger: The Analysis of Gene Expression Data: Methods and Software.
Salsburg: The Use of Restricted Significance Tests in Clinical Trials.
Simon/Korn/McShane/Radmacher/Wright/Zhao: Design and Analysis of DNA Microarray Investigations.
Sorensen/Gianola: Likelihood, Bayesian, and MCMC Methods in Quantitative Genetics.
Therneau/Grambsch: Modeling Survival Data: Extending the Cox Model.
Zhang/Singer: Recursive Partitioning in the Health Sciences.

SURVIVAL ANALYSIS
Techniques for Censored and Truncated Data

Second Edition

John P. Klein
Medical College of Wisconsin

Melvin L. Moeschberger
The Ohio State University Medical Center

With 97 Illustrations

Springer

clude those available in most standard computer packages. Techniques for assessing the fit of these parametric models are also discussed.

The final theme is multivariate models for survival data. In Chapter 13, tests for association between event times, adjusted for covariates, are given. An introduction to estimation in a frailty or random effect model is presented. An alternative approach to adjusting for association between some individuals based on an analysis of an independent working model is also discussed.

There should be ample material in this book for a one or two semester course for graduate students. A basic one semester or one quarter course would cover the following sections:

Chapter 2
Chapter 3, Sections 1–5
Chapter 4
Chapter 7, Sections 1–6, 8
Chapter 8
Chapter 9, Sections 1–4
Chapter 11
Chapter 12

In such a course the outlines of theoretical development of the techniques, in the theoretical notes, would be omitted. Depending on the length of the course and the interest of the instructor, these details could be added if the material in section 3.6 were covered or additional topics from the remaining chapters could be added to this skeleton outline. Applied exercises are provided at the end of the chapters. Solutions to odd numbered exercises are new to the second edition. The data used in the examples and in most of the exercises is available from us at our Web site which is accessible through the Springer Web site at http://www.springer-ny.com or http://www.biostat.mcw.edu/homepgs/klein/book.html.

Milwaukee, Wisconsin — John P. Klein
Columbus, Ohio — Melvin L. Moeschberger

Contents

1 Examples of Survival Data

1.1 Introduction

The problem of analyzing time to event data arises in a number of applied fields, such as medicine, biology, public health, epidemiology, engineering, economics, and demography. Although the statistical tools we shall present are applicable to all these disciplines, our focus is on applying the techniques to biology and medicine. In this chapter, we present some examples drawn from these fields that are used throughout the text to illustrate the statistical techniques we shall describe.

A common feature of these data sets is they contain either *censored* or *truncated observations*. Censored data arises when an individual's life length is known to occur only in a certain period of time. Possible censoring schemes are *right censoring*, where all that is known is that the individual is still alive at a given time, *left censoring* when all that is known is that the individual has experienced the event of interest prior to the start of the study, or *interval censoring*, where the only information is that the event occurs within some interval. Truncation schemes are *left truncation*, where only individuals who survive a sufficient time are included in the sample and *right truncation*, where only individuals who have experienced the event by a specified time are included in the sample. The issues of censoring and truncation are defined more carefully in Chapter 3.

1.2 Remission Duration from a Clinical Trial for Acute Leukemia

Freireich et al. (1963) report the results of a clinical trial of a drug 6-mercaptopurine (6-MP) versus a placebo in 42 children with acute leukemia. The trial was conducted at 11 American hospitals. Patients were selected who had a complete or partial remission of their leukemia induced by treatment with the drug prednisone. (A complete or partial remission means that either most or all signs of disease had disappeared from the bone marrow.) The trial was conducted by matching pairs of patients at a given hospital by remission status (complete or partial) and randomizing within the pair to either a 6-MP or placebo maintenance therapy. Patients were followed until their leukemia returned (relapse) or until the end of the study (in months). The data is reported in Table 1.1.

TABLE 1.1
Remission duration of 6-MP versus placebo in children with acute leukemia

Pair	*Remission Status at Randomization*	*Time to Relapse for Placebo Patients*	*Time to Relapse for 6-MP Patients*
1	Partial Remission	1	10
2	Complete Remission	22	7
3	Complete Remission	3	32+
4	Complete Remission	12	23
5	Complete Remission	8	22
6	Partial Remission	17	6
7	Complete Remission	2	16
8	Complete Remission	11	34+
9	Complete Remission	8	32+
10	Complete Remission	12	25+
11	Complete Remission	2	11+
12	Partial Remission	5	20+
13	Complete Remission	4	19+
14	Complete Remission	15	6
15	Complete Remission	8	17+
16	Partial Remission	23	35+
17	Partial Remission	5	6
18	Complete Remission	11	13
19	Complete Remission	4	9+
20	Complete Remission	1	6+
21	Complete Remission	8	10+

+Censored observation

This data set is used in Chapter 4 to illustrate the calculation of the estimated probability of survival using the product-limit estimator, the calculation of the Nelson-Aalen estimator of the cumulative hazard function, and the calculation of the mean survival time, along with their standard errors. It is further used in section 6.4 to estimate the survival function using Bayesian approaches. Matched pairs tests for differences in treatment efficacy are performed using the stratified log rank test in section 7.5 and the stratified proportional hazards model in section 9.3.

1.3 Bone Marrow Transplantation for Leukemia

Bone marrow transplants are a standard treatment for acute leukemia. Recovery following bone marrow transplantation is a complex process. Prognosis for recovery may depend on risk factors known at the time of transplantation, such as patient and/or donor age and sex, the stage of initial disease, the time from diagnosis to transplantation, etc. The final prognosis may change as the patient's posttransplantation history develops with the occurrence of events at random times during the recovery process, such as development of acute or chronic graft-versus-host disease (GVHD), return of the platelet count to normal levels, return of granulocytes to normal levels, or development of infections. Transplantation can be considered a failure when a patient's leukemia returns (relapse) or when he or she dies while in remission (treatment related death).

Figure 1.1 shows a simplified diagram of a patient's recovery process based on two intermediate events that may occur in the recovery process. These intermediate events are the possible development of acute GVHD that typically occurs within the first 100 days following transplantation and the recovery of the platelet count to a self-sustaining level $\geq 40 \times 10^9/l$ (called platelet recovery in the sequel). Immediately following transplantation, patients have depressed platelet counts and are free of acute GVHD. At some point, they may develop acute GVHD or have their platelets recover at which time their prognosis (probabilities of treatment related death or relapse at some future time) may change. These events may occur in any order, or a patient may die or relapse without any of these events occurring. Patients may, then, experience the other event, which again modifies their prognosis, or they may die or relapse.

To illustrate this process we consider a multicenter trial of patients prepared for transplantation with a radiation-free conditioning regimen.

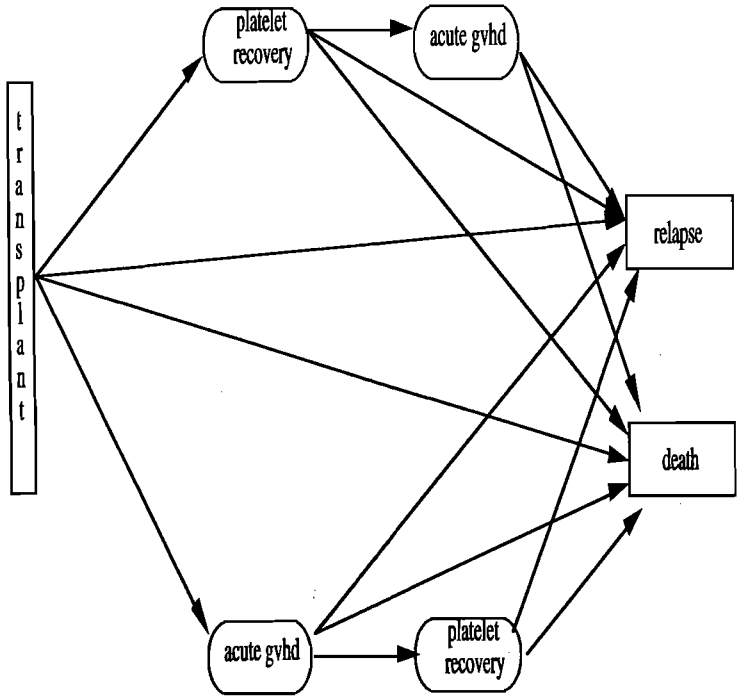

Figure 1.1 *Recovery Process from a Bone Marrow Transplant*

Details of the study are found in Copelan et al. (1991). The preparative regimen used in this study of allogeneic marrow transplants for patients with acute myeloctic leukemia (AML) and acute lymphoblastic leukemia (ALL) was a combination of 16 mg/kg of oral Busulfan (BU) and 120 mg/kg of intravenous cyclophosphamide (Cy). A total of 137 patients (99 AML, 38 ALL) were treated at one of four hospitals: 76 at The Ohio State University Hospitals (OSU) in Columbus; 21 at Hahnemann University (HU) in Philadelphia; 23 at St. Vincent's Hospital (SVH) in Sydney Australia; and 17 at Alfred Hospital (AH) in Melbourne. The study consists of transplants conducted at these institutions from March 1, 1984, to June 30, 1989. The maximum follow-up was 7 years. There were 42 patients who relapsed and 41 who died while in remission. Twenty-six patients had an episode of acute GVHD, and 17 patients

either relapsed or died in remission without their platelets returning to normal levels.

Several potential risk factors were measured at the time of transplantation. For each disease, patients were grouped into risk categories based on their status at the time of transplantation. These categories were as follows: ALL (38 patients), AML low-risk first remission (54 patients), and AML high-risk second remission or untreated first relapse (15 patients) or second or greater relapse or never in remission (30 patients). Other risk factors measured at the time of transplantation included recipient and donor gender (80 and 88 males respectively), recipient and donor cytomegalovirus immune status (CMV) status (68 and 58 positive, respectively), recipient and donor age (ranges 7–52 and 2–56, respectively), waiting time from diagnosis to transplantation (range 0.8–87.2 months, mean 19.7 months), and, for AML patients, their French-American-British (FAB) classification based on standard morphological criteria. AML patients with an FAB classification of M4 or M5 (45/99 patients) were considered to have a possible elevated risk of relapse or treatment-related death. Finally, patients at the two hospitals in Australia (SVH and AH) were given a graft-versus-host prophylactic combining methotrexate (MTX) with cyclosporine and possibly methylprednisolone. Patients at the other hospitals were not given methotrexate but rather a combination of cyclosporine and methylprednisolone. The data is presented in Table D.1 of Appendix D.

This data set is used throughout the book to illustrate the methods presented. In Chapter 4, it is used to illustrate the product-limit estimator of the survival function and the Nelson–Aalen estimator of the cumulative hazard rate of treatment failure. Based on these statistics, pointwise confidence intervals and confidence bands for the survival function are constructed. The data is also used to illustrate point and interval estimation of summary survival parameters, such as the mean and median time to treatment failure in this chapter.

This data set is also used in Chapter 4 to illustrate summary probabilities for competing risks. The competing risks, where the occurrence of one event precludes the occurrence of the other event, in this example, are relapse and death.

In section 6.2, the data set is used to illustrate the construction of estimates of the hazard rate. These estimates are based on smoothing the crude estimates of the hazard rate obtained from the jumps of the Nelson–Aalen estimator found in Chapter 4 using a weighted average of these estimates in a small interval about the time of interest. The weights are chosen using a kernel weighting function.

In Chapter 7, this data is used to illustrate tests for the equality of K survival curves. Both stratified and unstratified tests are discussed.

In Chapter 8, the data is used to illustrate tests of the equality of K hazard rates adjusted for possible fixed-time confounders. A proportional hazards model is used to make this adjustment. Model building for this problem is illustrated. In Chapter 9, the models found in Chap-

ter 8 are further refined to include covariates, whose values change over time, and to allow for stratified regression models. In Chapter 11, regression diagnostics for these models are presented.

1.4 Times to Infection of Kidney Dialysis Patients

In a study (Nahman et al., 1992) designed to assess the time to first exit-site infection (in months) in patients with renal insufficiency, 43 patients utilized a surgically placed catheter (Group 1), and 76 patients utilized a percutaneous placement of their catheter (Group 2). Cutaneous exit-site infection was defined as a painful cutaneous exit site and positive cultures, or peritonitis, defined as a presence of clinical symptoms, elevated peritoneal dialytic fluid, elevated white blood cell count (100 white blood cells/μl with >50% neutrophils), and positive peritoneal dialytic fluid cultures. The data appears in Table 1.2.

TABLE 1.2
Times to infection (in months) of kidney dialysis patients with different catheterization procedures

Surgically Placed Catheter
Infection Times: 1.5, 3.5, 4.5, 4.5, 5.5, 8.5, 8.5, 9.5, 10.5, 11.5, 15.5, 16.5, 18.5, 23.5, 26.5
Censored Observations: 2.5, 2.5, 3.5, 3.5, 3.5, 4.5, 5.5, 6.5, 6.5, 7.5, 7.5, 7.5, 7.5, 8.5, 9.5, 10.5, 11.5, 12.5, 12.5, 13.5, 14.5, 14.5, 21.5, 21.5, 22.5, 22.5, 25.5, 27.5
Percutaneous Placed Catheter
Infection Times: 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 2.5, 2.5, 3.5, 6.5, 15.5
Censored Observations: 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 1.5, 1.5, 1.5, 1.5, 2.5, 2.5, 2.5, 2.5, 2.5, 3.5, 3.5, 3.5, 3.5, 3.5, 4.5, 4.5, 4.5, 5.5, 5.5, 5.5, 5.5, 5.5, 6.5, 7.5, 7.5, 7.5, 8.5, 8.5, 8.5, 9.5, 9.5, 10.5, 10.5, 10.5, 11.5, 11.5, 12.5, 12.5, 12.5, 12.5, 14.5, 14.5, 16.5, 16.5, 18.5, 19.5, 19.5, 19.5, 20.5, 22.5, 24.5, 25.5, 26.5, 26.5, 28.5

The data is used in section 7.3 to illustrate how the inference about the equality of two survival curves, based on a two-sample weighted, log-rank test, depends on the choice of the weight function. In section 7.7, it is used to illustrate the two-sample Cramer–von Mises test for censored data. In the context of the proportional hazards model, this data is used in Chapter 8 to illustrate the different methods of constructing the partial likelihoods and the subsequent testing of equality of the survival

curves when there are ties present. Testing for proportional hazards is illustrated in section 9.2. The test reveals that a proportional hazards assumption for this data is not correct. A model with a time-varying, covariate effect is more appropriate, and in that section the optimal cutoff for "early" and "late" covariate effect on survival is found.

1.5 Times to Death for a Breast-Cancer Trial

In a study (Sedmak et al., 1989) designed to determine if female breast cancer patients, originally classified as lymph node negative by standard light microscopy (SLM), could be more accurately classified by immunohistochemical (IH) examination of their lymph nodes with an anticytokeratin monoclonal antibody cocktail, identical sections of lymph nodes were sequentially examined by SLM and IH. The significance of this study is that 16% of patients with negative axillary lymph nodes, by standard pathological examination, develop recurrent disease within 10 years. Forty-five female breast-cancer patients with negative axillary lymph nodes and a minimum 10-year follow-up were selected from The Ohio State University Hospitals Cancer Registry. Of the 45 patients, 9 were immunoperoxidase positive, and the remaining 36 remained negative. Survival times (in months) for both groups of patients are given in Table 1.3 ($^+$ denotes a censored observation).

TABLE 1.3
Times to death (in months) for breast cancer patients with different immunohistochemical responses

Immunoperoxidase Negative: 19, 25, 30, 34, 37, 46, 47, 51, 56, 57, 61, 66, 67, 74, 78, 86, 122^+, 123^+, 130^+, 130^+, 133^+, 134^+, 136^+, 141^+, 143^+ ,148^+, 151^+, 152^+,153^+,154^+, 156^+, 162^+, 164^+, 165^+, 182^+,189^+,
Immunoperoxidase Positive: 22, 23, 38, 42, 73, 77, 89, 115, 144^+

$^+$Censored observation

This data is used to show the construction of the likelihood function and in calculating a two-sample test based on proportional hazards with no ties with right-censored data in Chapter 8. It is also used in Chapter 10 to illustrate the least-squares estimation methodology in the context of the additive hazards model. In that chapter, we also used this data to illustrate estimation for an additive model with constant excess risk over time.

1.6 Times to Infection for Burn Patients

In a study (Ichida et al., 1993) to evaluate a protocol change in disinfectant practices in a large midwestern university medical center, 154 patient records and charts were reviewed. Infection of a burn wound is a common complication resulting in extended hospital stays and in the death of severely burned patients. Control of infection remains a prominent component of burn management. The purpose of this study is to compare a routine bathing care method (initial surface decontamination with 10% povidone-iodine followed with regular bathing with Dial soap) with a body-cleansing method using 4% chlorhexidine gluconate. Medical records of patients treated during the 18-month study period provided information on burn wound infections and other medical information. The time until staphylococcus infection was recorded (in days), along with an indicator variable—whether or not an infection had occurred. Other fixed covariates recorded were gender (22% female), race (88% white), severity of the burn as measured by percentage of total surface area of body burned (average of 24.7% range 2–95%), burn site (45% on head, 23% on buttocks, 84% on trunk, 41% on upper legs, 31% on lower legs, or 29% in respiratory tract), and type of burn (6% chemical, 12% scald, 7% electric, or 75% flame). Two time-dependent covariates were recorded, namely, time to excision and time to prophylactic antibiotic treatment administered, along with the two corresponding indicator variables, namely, whether the patient's wound had been excised (64%) and whether the patient had been treated with an antibiotic sometime during the course of the study (41%). Eighty-four patients were in the group which received the new bathing solution, chlorhexidine, and 70 patients served as the historical control group which received the routine bathing care, povidone-iodine. The data is available on the authors' web site and is used in the exercises.

1.7 Death Times of Kidney Transplant Patients

Data on the time to death of 863 kidney transplant patients is available on the authors' web site. All patients had their transplant performed at The Ohio State University Transplant Center during the period 1982–1992. The maximum follow-up time for this study was 9.47 years. Patients were censored if they moved from Columbus (lost to follow-up) or if they were alive on June 30, 1992. In the sample, there were 432 white males, 92 black males, 280 white females, and 59 black females. Patient ages at transplant ranged from 9.5 months to 74.5 years with a mean age of 42.8 years. Seventy-three (16.9%) of the white males, 14 (15.2%) of the black males, 39 (13.9%) of the white females and 14 (23.7%) of the black females died prior to the end of the study.

In Chapter 6, the problem of estimating the hazard rate, using a kernel smoothing procedure, is discussed. In particular, the effect of changing the bandwidth and the choice of a kernel are considered. In Chapter 8 this data is also used to illustrate methods for discretizing a continuous covariate.

1.8 Death Times of Male Laryngeal Cancer Patients

Kardaun (1983) reports data on 90 males diagnosed with cancer of the larynx during the period 1970–1978 at a Dutch hospital. Times recorded are the intervals (in years) between first treatment and either death or the end of the study (January 1, 1983). Also recorded are the patient's age at the time of diagnosis, the year of diagnosis, and the stage of the patient's cancer. The four stages of disease in the study were based on the T.N.M. (primary tumor (T), nodal involvement (N) and distant metastasis (M) grading) classification used by the American Joint Committee for Cancer Staging (1972). The four groups are Stage I, $T_1N_0M_0$ with 33 patients; Stage II, $T_2N_0M_0$ with 17 patients; Stage III, $T_3N_0M_0$ and $T_xN_1M_0$, with 27 patients; $x = 1$, 2, or 3; and Stage IV, all other TNM combinations except TIS with 13 patients. The stages are ordered from least serious to most serious. The data is available on the authors' web site.

In section 7.4, the data is used to illustrate a test for trend to confirm the hypothesis that the higher the stage the greater the chance of dying. In Chapter 8, a global test for the effect of stage on survival is performed in the context of the proportional hazards model, and local tests are illustrated, after an adjustment for the patient's age. An analysis of variance (ANOVA) table is presented to summarize the effects of stage and age on survival. Contrasts are used to test the hypothesis that linear combinations of stage effects are zero. The construction of confidence intervals for different linear combinations of the stage effects is illustrated. The concept of an interaction in a proportional hazards regression model is illustrated through a stage by age interaction factor. The survival curve is estimated for each stage based on the Cox proportional hazards model.

This data is also used in Chapter 10 to illustrate estimation methodology in the additive hazards model. In Chapter 12, this data set is used to illustrate the fit of parametric models, using the accelerated failure-time model. The goodness of fit of these models is also discussed. The log logistic model is used in section 12.5 to illustrate using deviance residuals.

1.9 Autologous and Allogeneic Bone Marrow Transplants

The data in Table 1.4 is a sample of 101 patients with advanced acute myelogenous leukemia reported to the International Bone Marrow Transplant Registry. Fifty-one of these patients had received an autologous (auto) bone marrow transplant in which, after high doses of chemotherapy, their own marrow was reinfused to replace their destroyed immune system. Fifty patients had an allogeneic (allo) bone marrow transplant where marrow from an HLA (Histocompatibility Leukocyte Antigen)-matched sibling was used to replenish their immune systems.

An important question in bone marrow transplantation is the comparison of the effectiveness of these two methods of transplant as measured by the length of patients' leukemia-free survival, the length of time they are alive, and how long they remain free of disease after their transplants. In Chapter 7, this comparison is made using a weighted log-rank test, and a censored data version of the median test and the t-test.

TABLE 1.4
Leukemia free-survival times (in months) for Autologous and Allogeneic Transplants

The leukemia-free survival times for the 50 allo transplant patients were 0.030, 0.493, 0.855, 1.184, 1.283, 1.480, 1.776, 2.138, 2.500, 2.763, 2.993, 3.224, 3.421, 4.178, 4.441$^+$, 5.691, 5.855$^+$, 6.941$^+$, 6.941, 7.993$^+$, 8.882, 8.882, 9.145$^+$, 11.480, 11.513, 12.105$^+$, 12.796, 12.993$^+$, 13.849$^+$, 16.612$^+$, 17.138$^+$, 20.066, 20.329$^+$, 22.368$^+$, 26.776$^+$, 28.717$^+$, 28.717$^+$, 32.928$^+$, 33.783$^+$, 34.211$^+$, 34.770$^+$, 39.539$^+$, 41.118$^+$, 45.033$^+$, 46.053$^+$, 46.941$^+$, 48.289$^+$, 57.401$^+$, 58.322$^+$, 60.625$^+$;
and, for the 51 auto patients, 0.658, 0.822, 1.414, 2.500, 3.322, 3.816, 4.737, 4.836$^+$, 4.934, 5.033, 5.757, 5.855, 5.987, 6.151, 6.217, 6.447$^+$, 8.651, 8.717, 9.441$^+$, 10.329, 11.480, 12.007, 12.007$^+$, 12.237, 12.401$^+$, 13.059$^+$, 14.474$^+$, 15.000$^+$, 15.461, 15.757, 16.480, 16.711, 17.204$^+$, 17.237, 17.303$^+$, 17.664$^+$, 18.092, 18.092$^+$, 18.750$^+$, 20.625$^+$, 23.158, 27.730$^+$, 31.184$^+$, 32.434$^+$, 35.921$^+$, 42.237$^+$, 44.638$^+$, 46.480$^+$, 47.467$^+$, 48.322$^+$, 56.086.

As usual, $^+$ denotes a censored observation.

This data is used in Chapter 11 to illustrate graphical methods for checking model assumptions following a proportional hazards regression analysis. In section 11.3, the martingale residuals are used to check overall model fit. In section 11.4, score residuals are used to check the proportional hazards assumption on disease-free survival for type of transplant. In section 11.5, the use of deviance residuals is illustrated for checking for outliers and, in section 11.6, the influence of individual observations is examined graphically.

In Chapter 12, this data set is used to illustrate the fit of parametric models using the accelerated failure-time model. The goodness of fit of these models is also discussed. Diagnostic plots for checking the fit of a parametric regression model using this data set are illustrated in section 12.5.

1.10 Bone Marrow Transplants for Hodgkin's and Non-Hodgkin's Lymphoma

The data in Table 1.5 was collected on 43 bone marrow transplant patients at The Ohio State University Bone Marrow Transplant Unit. Details of this study can be found in Avalos et al. (1993). All patients had either Hodgkin's disease (HOD) or non-Hodgkin's lymphoma (NHL) and were given either an allogeneic (Allo) transplant from an HLA match sibling donor or an autogeneic (Auto) transplant; i.e., their own marrow was cleansed and returned to them after a high dose of chemotherapy. Also included are two possible explanatory variables, Karnofsky score at transplant and the waiting time in months from diagnosis to transplant. Of interest is a test of the null hypothesis of no difference in the leukemia-free survival rate between patients given an Allo or Auto transplant, adjusting for the patient's disease state. This test, which requires stratification of the patient's disease, is presented in section 7.5. We also use this data in section 11.3 to illustrate how the martingale residual can be used to determine the functional form of a covariate. The data, in Table 1.5, consists of the time on study for each patient, T_i, and the event indicator $\delta_i = 1$ if dead or relapsed; 0 otherwise; and two covariates Z_1, the pretransplant Karnofsky score and Z_2, the waiting time to transplant.

TABLE 1.5
Times to death or relapse (in days) for patients with bone marrow transplants for Hodgkin's and non-Hodgkin's lymphoma

Allo NHL				*Auto NHL*				*Allo HOD*				*Auto HOD*			
T_i	δ_i	Z_1	Z_2	T_i	δ_i	Z_1	Z_2	T_i	δ_i	Z_1	Z_2	T_i	δ_i	Z_1	Z_2
28	1	90	24	42	1	80	19	2	1	20	34	30	1	90	73
32	1	30	7	53	1	90	17	4	1	50	28	36	1	80	61
49	1	40	8	57	1	30	9	72	1	80	59	41	1	70	34
84	1	60	10	63	1	60	13	77	1	60	102	52	1	60	18
357	1	70	42	81	1	50	12	79	1	70	71	62	1	90	40
933	0	90	9	140	1	100	11					108	1	70	65
1078	0	100	16	176	1	80	38					132	1	60	17
1183	0	90	16	210	0	90	16					180	0	100	61
1560	0	80	20	252	1	90	21					307	0	100	24
2114	0	80	27	476	0	90	24					406	0	100	48
2144	0	90	5	524	1	90	39					446	0	100	52
				1037	0	90	84					484	0	90	84
												748	0	90	171
												1290	0	90	20
												1345	0	80	98

TABLE 1.6
Death times (in weeks) of patients with cancer of the tongue

Aneuploid Tumors:
Death Times: 1, 3, 3, 4, 10, 13, 13, 16, 16, 24, 26, 27, 28, 30, 30, 32, 41, 51, 65, 67, 70, 72, 73, 77, 91, 93, 96, 100, 104, 157, 167
Censored Observations: 61, 74, 79, 80, 81, 87, 87, 88, 89, 93, 97, 101, 104, 108, 109, 120, 131, 150, 231, 240, 400
Diploid Tumors:
Death Times: 1, 3, 4, 5, 5, 8, 12, 13, 18, 23, 26, 27, 30, 42, 56, 62, 69, 104, 104, 112, 129, 181
Censored Observations: 8, 67, 76, 104, 176, 231

1.11 Times to Death for Patients with Cancer of the Tongue

A study was conducted on the effects of ploidy on the prognosis of patients with cancers of the mouth. Patients were selected who had a paraffin-embedded sample of the cancerous tissue taken at the time of surgery. Follow-up survival data was obtained on each patient. The tissue samples were examined using a flow cytometer to determine if the tumor had an aneuploid (abnormal) or diploid (normal) DNA profile using a technique discussed in Sickle–Santanello et al. (1988). The data in Table 1.6 is on patients with cancer of the tongue. Times are in weeks.

The data is used in exercises.

1.12 Times to Reinfection for Patients with Sexually Transmitted Diseases

A major problem in certain subpopulations is the occurrence of sexually transmitted diseases (STD). Even if one ignores the lethal effects of the Acquired Immune Deficiency Syndrome (AIDS), other sexually transmitted diseases still have a significant impact on the morbidity of the community. Two of these sexually transmitted diseases are the focus of this investigation: gonorrhea and chlamydia. These diseases are of special interest because they are often asymptomatic in the female, and, if left untreated, can lead to complications including sterility.

Both of these diseases can be easily prevented and effectively treated. Therefore, it is a mystery why the incidence of these diseases remain high in several subpopulations. One theory is that a core group of individuals experience reinfections, thereby, serving as a natural reservoir of the disease organism and transferring the disease to uninfected individuals.

The purpose of this study is to identify those factors which are related to time until reinfection by either gonorrhea or chlamydia, given an initial infection of gonorrhea or chlamydia. A sample of 877 individuals, with an initial diagnosis of gonorrhea or chlamydia were followed for reinfection. In addition to the primary outcome variable just stated, an indicator variable which indicates whether a reinfection occurred was recorded. Demographic variables recorded were race (33% white, 67% black), marital status (7% divorced/separated, 3% married and 90% single), age of patient at initial infection (average age is 20.6 years with a range of 13–48 years), years of schooling (11.4 years with a range of 6–18 years), and type of initial infection (16% gonorrhea, 45% chlamydia and 39% both gonorrhea and chlamydia). Behavioral factors recorded at the examination, when the initial infection was diagnosed, were number of partners in the last 30 days (average is 1.27 with a range of 0–19), oral sex within past 12 months (33%), rectal sex within past 12 months (6%), and condom use (6% always, 58% sometimes, and 36%

never). Symptoms noticed at time of initial infection were presence of abdominal pain (14%), sign of discharge (46%), sign of dysuria (13%), sign of itch (19%), sign of lesion (3%), sign of rash (3%), and sign of lymph involvement (1%). If the factors related to a greater risk of reinfection can be identified, then, interventions could be targeted to those individuals who are at greatest risk for reinfection. This, in turn, should reduce the size of the core group and, thereby, reduce the incidence of the diseases. The data for this study is available on our web site.

This data is used in the exercises.

1.13 Time to Hospitalized Pneumonia in Young Children

Data gathered from 3,470 annual personal interviews conducted for the National Longitudinal Survey of Youth (NLSY, 1995) from 1979 through 1986 was used to study whether the mother's feeding choice (breast feeding vs. never breast fed) protected the infant against hospitalized pneumonia in the first year of life. Information obtained about the child included whether it had a normal birthweight, as defined by weighing at least 5.5 pounds (36%), race (56% white, 28% black, and 16% other), number of siblings (range 0–6), and age at which the child was hospitalized for pneumonia, along with an indicator variable as to whether the child was hospitalized. Demographic characteristics of the mother, such as age (average is 21.64 years with a range of 14–29 years), years of schooling (average of 11.4 years), region of the country (15% Northeast, 25% North central, 40% South, and 20% West), poverty (92%), and urban environment (76%). Health behavior measures during pregnancy, such as alcohol use (36%) and cigarette use (34%), were also recorded. The data for this study is available on our web site.

This data is used in the exercises.

1.14 Times to Weaning of Breast-Fed Newborns

The National Longitudinal Survey of Youth is a stratified random sample which was begun in 1979. Youths, aged 14 to 21 in 1979, have been interviewed yearly through 1988. Beginning in 1983, females in the survey were asked about any pregnancies that have occurred since they

were last interviewed (pregnancies before 1983 were also documented). Questions regarding breast feeding are included in the questionnaire.

This data set consists of the information from 927 first-born children to mothers who chose to breast feed their children and who have complete information for all the variables of interest. The sample was restricted to children born after 1978 and whose gestation age was between 20 and 45 weeks. The year of birth restriction was included in an attempt to eliminate recall problems.

The response variable in the data set is duration of breast feeding in weeks, followed by an indicator of whether the breast feeding was completed (i.e., the infant is weaned). Explanatory variables for breast-feeding duration include race of mother (1 if white, 2 if black, 3 if other); poverty status indicator (1 if mother in poverty); smoking status of mother (1 if smoking at birth of child); alcohol-drinking status of mother (1 if drinking at birth of child); age of mother at child's birth, year of child's birth, education of mother (in years); and lack of prenatal care status (1 if mother sought prenatal care after third month or never sought prenatal care, 0 if mother sought prenatal care in first three months of pregnancy). The complete data for this study is available on our web site.

This data is used in section 5.4 to illustrate the construction of the cohort life table. In Chapter 8, it is used to show how to build a model where predicting the outcome is the main purpose, i.e., interest is in finding factors which contribute to the distribution of the time to weaning.

1.15 Death Times of Psychiatric Patients

Woolson (1981) has reported survival data on 26 psychiatric inpatients admitted to the University of Iowa hospitals during the years 1935–1948. This sample is part of a larger study of psychiatric inpatients discussed by Tsuang and Woolson (1977). Data for each patient consists of age at first admission to the hospital, sex, number of years of follow-up (years from admission to death or censoring) and patient status at the follow-up time. The data is given in Table 1.7. In section 6.3, the estimate of the relative mortality function and cumulative excess mortality of these patients, compared to the standard mortality rates of residents of Iowa in 1959, is considered. In section 7.2, this data is used to illustrate one-sample hypothesis tests. Here, a comparison of the survival experience of these 26 patients is made to the standard mortality of residents of Iowa to determine if psychiatric patients tend to have shorter lifetimes. It is used in Chapter 9 to illustrate left truncation in the context of proportional hazards models.

TABLE 1.7
Survival data for psychiatric inpatients

Gender	*Age at Admission*	*Time of Follow-up*
Female	51	1
Female	58	1
Female	55	2
Female	28	22
Male	21	30+
Male	19	28
Female	25	32
Female	48	11
Female	47	14
Female	25	36+
Female	31	31+
Male	24	33+
Male	25	33+
Female	30	37+
Female	33	35+
Male	36	25
Male	30	31+
Male	41	22
Female	43	26
Female	45	24
Female	35	35+
Male	29	34+
Male	35	30+
Male	32	35
Female	36	40
Male	32	39+

+Censored observation

1.16 Death Times of Elderly Residents of a Retirement Community

Channing House is a retirement center located in Palo Alto, California. Data on ages at death of 462 individuals (97 males and 365 females) who were in residence during the period January 1964 to July 1975 has been reported by Hyde (1980). A distinctive feature of these individuals was that all were covered by a health care program provided by the center which allowed for easy access to medical care without any additional financial burden to the resident. The age in months when members of the community died or left the center and the age when individuals entered the community is available on the authors' web site.

The life lengths in this data set are *left truncated* because an individual must survive to a sufficient age to enter the retirement community. Individuals who die at an early age are excluded from the study. Ignoring this left truncation leads to the problem of length-biased sampling. The concept of left truncation and the bias induced into the estimation process by ignoring it is discussed in section 3.4.

This data will be used in section 4.6 to illustrate how one estimates the conditional survival function for left-truncated data. The data is used in section 7.3 to illustrate the comparison of two samples (male and female), when there is left truncation and right censoring employing the log-rank test, and in Chapter 9 employing the Cox proportional hazards model.

1.17 Time to First Use of Marijuana

Turnbull and Weiss (1978) report part of a study conducted at the Stanford-Palo Alto Peer Counseling Program (see Hamburg et al. [1975] for details of the study). In this study, 191 California high school boys were asked, "When did you first use marijuana?" The answers were the exact ages (uncensored observations); "I never used it," which are right-censored observations at the boys' current ages; or "I have used it but can not recall just when the first time was," which is a left-censored observation (see section 3.3). Notice that a left-censored observation

TABLE 1.8
Marijuana use in high school boys

Age	*Number of Exact Observations*	*Number Who Have Yet to Smoke Marijuana*	*Number Who Have Started Smoking at an Earlier Age*
10	4	0	0
11	12	0	0
12	19	2	0
13	24	15	1
14	20	24	2
15	13	18	3
16	3	14	2
17	1	6	3
18	0	0	1
>18	4	0	0

tells us only that the event has occurred prior to the boy's current age. The data is in Table 1.8.

This data is used in section 5.2 to illustrate the calculation of the survival function for both left- and right-censored data, commonly referred to as doubly censored data.

1.18 Time to Cosmetic Deterioration of Breast Cancer Patients

Beadle et al. (1984a and b) report a retrospective study carried out to compare the cosmetic effects of radiotherapy alone versus radiotherapy and adjuvant chemotherapy on women with early breast cancer. The use of an excision biopsy, followed by radiation therapy, has been suggested as an alternative to mastectomy. This therapy preserves the breast and, hence, has the benefit of an improved cosmetic effect. The use of adjuvant chemotherapy is often indicated to prevent recurrence of the cancer, but there is some clinical evidence that it enhances the effects of radiation on normal tissue, thus, offsetting the cosmetic benefit of this procedure.

To compare the two treatment regimes, a retrospective study of 46 radiation only and 48 radiation plus chemotherapy patients was made. Patients were observed initially every 4–6 months, but, as their recovery progressed, the interval between visits lengthened. At each visit, the clinician recorded a measure of breast retraction on a three-point scale (none, moderate, severe). The event of interest was the time to first

TABLE 1.9

Time to cosmetic deterioration (in months) in breast cancer patients with two treatment regimens

Radiotherapy only: (0, 7]; (0, 8]; (0, 5]; (4, 11]; (5, 12]; (5, 11]; (6, 10]; (7, 16]; (7, 14]; (11, 15]; (11, 18]; ≥15; ≥17; (17, 25]; (17, 25]; ≥18; (19, 35]; (18, 26]; ≥22; ≥24; ≥24; (25, 37]; (26, 40]; (27, 34]; ≥32; ≥33; ≥34; (36, 44]; (36, 48]; ≥36; ≥36; (37, 44]; ≥37; ≥37; ≥37; ≥38; ≥40; ≥45; ≥46; ≥46; ≥46; ≥46; ≥46; ≥46; ≥46; ≥46.

Radiotherapy and Chemotherapy: (0, 22]; (0, 5]; (4, 9]; (4, 8]; (5, 8]; (8, 12]; (8, 21]; (10, 35]; (10, 17]; (11, 13]; ≥11; (11, 17]; ≥11; (11, 20]; (12, 20]; ≥13; (13, 39]; ≥13; ≥13; (14, 17]; (14, 19]; (15, 22]; (16, 24]; (16, 20]; (16, 24]; (16, 60]; (17, 27]; (17, 23]; (17, 26]; (18, 25]; (18, 24]; (19, 32]; ≥21; (22, 32]; ≥23; (24, 31]; (24, 30]; (30, 34]; (30, 36]; ≥31; ≥32; (33, 40]; ≥34; ≥34; ≥35; (35, 39]; (44, 48]; ≥48.

(a, b]—interval in which deterioration took place.

appearance of moderate or severe breast retraction. Due to the fact that patients were observed only at these random times, the exact time of breast retraction is known only to fall in the interval between visits. This type of data is call *interval-censored data* (see section 3.3).

The data for the two groups is shown in Table 1.9. The data consists of the interval, in months, in which deterioration occurred or the last time a patient was seen without deterioration having yet occurred (right-censored observations). This data is used in section 5.2 to illustrate the computation of an estimate of the survival function based on interval-censored data.

1.19 Time to AIDS

Lagakos et al. (1988) report data on the infection and induction times for 258 adults and 37 children who were infected with the AIDS virus and developed AIDS by June 30, 1986. The data consists of the time in years, measured from April 1, 1978, when adults were infected by the virus from a contaminated blood transfusion, and the waiting time to development of AIDS, measured from the date of infection. For the pediatric population, children were infected in utero or at birth, and the infection time is the number of years from April 1, 1978 to birth. The data is in Table 1.10.

In this sampling scheme, only individuals who have developed AIDS prior to the end of the study period are included in the study. Infected individuals who have yet to develop AIDS are not included in the sample. This type of data is called *right-truncated* data (see section 3.4). Estimation of the survival function for this data with right-truncated data is discussed in section 5.3.

TABLE 1.10
Induction times (in years) for AIDS in adults and children

Infection Time	*Adult Induction Time*	*Child Induction Time*
0.00	5	
0.25	6.75	
0.75	5, 5, 7.25	
1.00	4.25, 5.75, 6.25, 6.5	5.5
1.25	4, 4.25, 4.75, 5.75	
1.50	2.75, 3.75, 5, 5.5, 6.5	2.25
1.75	2.75, 3, 5.25, 5.25	
2.00	2.25, 3, 4, 4.5, 4.75, 5, 5.25, 5.25, 5.5, 5.5, 6	
2.25	3, 5.5	3
2.50	2.25, 2.25, 2.25, 2.25, 2.5, 2.75, 3, 3.25, 3.25, 4, 4, 4	
2.75	1.25, 1.5, 2.5, 3, 3, 3.25, 3.75, 4.5, 4.5, 5, 5, 5.25, 5.25, 5.25, 5.25, 5.25	1
3.00	2, 3.25, 3.5, 3.75, 4, 4, 4.25, 4.25, 4.25, 4.75, 4.75, 4.75, 5	1.75
3.25	1.25, 1.75, 2, 2, 2.75, 3, 3, 3.5, 3.5, 4.25, 4.5	
3.50	1.25, 2.25, 2.25, 2.5, 2.75, 2.75, 3, 3.25, 3.5, 3.5, 4, 4, 4.25, 4.5, 4.5	0.75
3.75	1.25, 1.75, 1.75, 2, 2.75, 3, 3, 3, 4, 4.25, 4.25	0.75, 1, 2.75, 3, 3.5, 4.25
4.00	1, 1.5, 1.5, 2, 2.25, 2.75, 3.5, 3.75, 3.75, 4	1
4.25	1.25, 1.5, 1.5, 2, 2, 2, 2.25, 2.5, 2.5, 2.5, 3, 3.5, 3.5	1.75
4.50	1, 1.5, 1.5, 1.5, 1.75, 2.25, 2.25, 2.5, 2.5, 2.5, 2.5, 2.75, 2.75, 2.75, 2.75, 3, 3, 3, 3.25, 3.25	3.25
4.75	1, 1.5, 1.5, 1.5, 1.75, 1.75, 2, 2.25, 2.75, 3, 3, 3.25, 3.25, 3.25, 3.25, 3.25, 3.25	1, 2.25
5.00	0.5, 1.5, 1.5, 1.75, 2, 2.25, 2.25, 2.25, 2.5, 2.5, 3, 3, 3	0.5, 0.75, 1.5, 2.5
5.25	0.25, 0.25, 0.75, 0.75, 0.75, 1, 1, 1.25, 1.25, 1.5, 1.5, 1.5, 1.5, 2.25, 2.25, 2.5, 2.5, 2.75	0.25, 1, 1.5
5.50	1, 1, 1, , 1.25, 1.25, 1.75, 2, 2.25, 2.25, 2.5	.5, 1.5, 2.5
5.75	0.25, 0.75, 1, 1.5, 1.5, 1.5, 2, 2, 2.25	1.75
600	0.5, 0.75, 0.75, 0.75, 1, 1, 1, 1.25, 1.25, 1.5, 1.5, 1.75, 1.75, 1.75, 2	0.5, 1.25
6.25	0.75, 1, 1.25, 1.75, 1.75	0.5, 1.25
6.50	0.25, 0.25, 0.75, 1, 1.25, 1.5	0.75
6.75	0.75, 0.75, 0.75, 1, 1.25, 1.25, 1.25	0.5, 0.75
7.00	0.75	0.75
7.25	0.25	0.25

2 Basic Quantities and Models

2.1 Introduction

In this chapter we consider the basic parameters used in modeling survival data. We shall define these quantities and show how they are interrelated in sections 2.2–2.4. In section 2.5 some common parametric models are discussed. The important application of regression to survival analysis is covered in section 2.6, where both parametric and semiparametric models are presented. Models for competing risks are discussed in section 2.7.

Let X be the time until some specified event. This event may be death, the appearance of a tumor, the development of some disease, recurrence of a disease, equipment breakdown, cessation of breast feeding, and so forth. Furthermore, the event may be a good event, such as remission after some treatment, conception, cessation of smoking, and so forth. More precisely, in this chapter, X is a nonnegative random variable from a homogeneous population. Four functions characterize the distribution of X, namely, the *survival function*, which is the probability of an individual surviving to time x; the *hazard rate (function)*, sometimes termed *risk function*, which is the chance an individual of age x experiences the event in the next instant in time; the *probability density (or probability mass) function*, which is the unconditional probability of the event's occurring at time x; and the *mean residual life* at time x, which is the mean time to the event of interest, given the event has not occurred at x. If we know any one of these four