4

Nonparametric Estimation of Basic Quantities for Right-Censored and Left-Truncated Data

4.1 Introduction

In this chapter we shall examine techniques for drawing an inference about the distribution of the time to some event X, based on a sample of right-censored survival data. A typical data point consists of a time on study and an indicator of whether this time is an event time or a censoring time for each of the n individuals in the study. We assume throughout this chapter that the potential censoring time is unrelated to the potential event time. The methods are appropriate for Type I, Type II, progressive or random censoring discussed in section 3.2.

To allow for possible ties in the data, suppose that the events occur at D distinct times $t_1 < t_2 < \cdots < t_D$, and that at time t_i there are d_i events (sometimes simply referred to as deaths). Let Y_i be the number of individuals who are at risk at time t_i. Note that Y_i is a count of the number

of individuals with a time on study of t_i or more (i.e., the number of individuals who are alive at t_i or experience the event of interest at t_i). The quantity d_i/Y_i provides an estimate of the conditional probability that an individual who survives to just prior to time t_i experiences the event at time t_i. As we shall see, this is the basic quantity from which we will construct estimators of the survival function and the cumulative hazard rate.

Basic estimators of the survival function $S(t)$, the cumulative hazard function $H(t)$, and their standard errors based on right-censored data are discussed in section 4.2. In section 4.3, confidence intervals for $S(t)$ and $H(t)$ for a fixed value of t are presented, and section 4.4 presents confidence bands which provide a specified coverage probability for a range of times. Section 4.5 discusses inference for the mean time to event and for percentiles of X based on right-censored data. The final section shows how the estimators developed for right-censored data can be extended to left-truncated data. Estimating for other censoring and truncating schemes is considered in Chapter 5.

4.2 Estimators of the Survival and Cumulative Hazard Functions for Right-Censored Data

The standard estimator of the survival function, proposed by Kaplan and Meier (1958), is called the Product-Limit estimator. This estimator is defined as follows for all values of t in the range where there is data:

$$\hat{S}(t) = \begin{cases} 1 & \text{if } t < t_1, \\ \prod_{t_i \le t}[1 - \frac{d_i}{Y_i}], & \text{if } t_1 \le t \end{cases}. \tag{4.2.1}$$

For values of t beyond the largest observation time this estimator is not well defined (see Practical Notes 1 and 2 for suggestions as to solutions to this problem).

The Product-Limit estimator is a step function with jumps at the observed event times. The size of these jumps depends not only on the number of events observed at each event time t_i, but also on the pattern of the censored observations prior to t_i.

The variance of the Product-Limit estimator is estimated by Greenwood's formula:

$$\hat{V}[\hat{S}(t)] = \hat{S}(t)^2 \sum_{t_i \le t} \frac{d_i}{Y_i(Y_i - d_i)}. \tag{4.2.2}$$

The standard error of the Product-Limit estimator is given by $\{\hat{V}[\hat{S}(t)]\}^{1/2}$.

EXAMPLE 4.1 We consider the data in section 1.2 on the time to relapse of patients in a clinical trial of 6-MP against a placebo. We shall consider only the 6-MP patients. The calculations needed to construct the Product-Limit estimator and its estimated variance are in Table 4.1A. The Product-Limit estimator, found in Table 4.1B, is a step function. Figure 4.1A shows a plot of this estimated survival function. Note that the survival curve is defined only up to 35 weeks, the largest of the observation times.

TABLE 4.1A

Construction of the Product-Limit Estimator and its Estimated Variance for the 6-MP Group

Time t_i	*Number of events* d_i	*Number at risk* Y_i	*Product-Limit Estimator* $\hat{S}(t) = \prod_{t_i \le t}[1 - \frac{d_i}{Y_i}]$	$\sum_{t_i \le t} \frac{d_i}{Y_i(Y_i - d_i)}$	$\hat{S}(t)^2 \sum_{t_i \le t} \frac{d_i}{Y_i(Y_i - d_i)}$.
6	3	21	$[1 - \frac{3}{21}] = 0.857$	$\frac{3}{21\times18} = 0.0079$	$0.857^2 \times 0.0079 = 0.0058$
7	1	17	$[0.857](1 - \frac{1}{17}) = 0.807$	$0.0079 + \frac{1}{17\times16} = 0.0116$	$0.807^2 \times 0.0116 = 0.0076$
10	1	15	$[0.807](1 - \frac{1}{15}) = 0.753$	$0.0116 + \frac{1}{15\times14} = 0.0164$	$0.753^2 \times 0.0164 = 0.0093$
13	1	12	$[0.753](1 - \frac{1}{12}) = 0.690$	$0.0164 + \frac{1}{12\times11} = 0.0240$	$0.690^2 \times 0.0240 = 0.0114$
16	1	11	$[0.690](1 - \frac{1}{11}) = 0.628$	$0.0240 + \frac{1}{11\times10} = 0.0330$	$0.628^2 \times 0.0330 = 0.0130$
22	1	7	$[0.628](1 - \frac{1}{7}) = 0.538$	$0.0330 + \frac{1}{7\times6} = 0.0569$	$0.538^2 \times 0.0569 = 0.0164$
23	1	6	$[0.538](1 - \frac{1}{6}) = 0.448$	$0.0569 + \frac{1}{6\times5} = 0.0902$	$0.448^2 \times 0.0902 = 0.0181$

TABLE 4.1B

The Product-Limit Estimator and Its Estimated Standard Error for the 6-MP Group

Time on Study (t)	$\hat{S}(t)$	*Standard Error*
$0 \le t < 6$	1.000	0.000
$6 \le t < 7$	0.857	0.076
$7 \le t < 10$	0.807	0.087
$10 \le t < 13$	0.753	0.096
$13 \le t < 16$	0.690	0.107
$16 \le t < 22$	0.628	0.114
$22 \le t < 23$	0.538	0.128
$23 \le t < 35$	0.448	0.135

The Product-Limit estimator provides an efficient means of estimating the survival function for right-censored data. It can also be used to estimate the cumulative hazard function $H(t) = -\ln[S(t)]$. The estimator is $\hat{H}(t) = -\ln[\hat{S}(t)]$. An alternate estimator of the cumulative

hazard rate, which has better small-sample-size performance than the estimator based on the Product-Limit estimator, was first suggested by Nelson (1972) in a reliability context. The estimator was rediscovered by Aalen (1978b) who derived the estimator using modern counting process techniques (see section 3.6 for a sketch of this derivation). This estimator, which shall be referred to as the Nelson–Aalen estimator of the cumulative hazard, is defined up to the largest observed time on study as follows:

$$\tilde{H}(t) = \begin{cases} 0, & \text{if } t \le t_1, \\ \sum_{t_i \le t} \frac{d_i}{Y_i}, & \text{if } t_1 \le t. \end{cases} \tag{4.2.3}$$

The estimated variance of the Nelson–Aalen estimator is due to Aalen (1978b) and is given by

$$\sigma_H^2(t) = \sum_{t_i \le t} \frac{d_i}{Y_i^2}. \tag{4.2.4}$$

Based on the Nelson–Aalen estimator of the cumulative hazard rate (4.2.3), an alternate estimator of the survival function is given by $\tilde{S}(t) = \exp[-\tilde{H}(t)]$.

The Nelson–Aalen estimator has two primary uses in analyzing data. The first is in selecting between parametric models for the time to

TABLE 4.2
Construction of the Nelson–Aalen Estimator and its Estimated Variance for the 6-MP Group

Time t	$\tilde{H}(t) = \sum_{t_i \le t} \frac{d_i}{Y_i}$	$\sigma_H^2 = \sum_{t_i \le t} \frac{d_i}{Y_i^2}$.	*Standard Error*
$0 \le t < 6$	0	0	0
$6 \le t < 7$	$\frac{3}{21} = 0.1428$	$\frac{3}{21^2} = 0.0068$	0.0825
$7 \le t < 10$	$0.1428 + \frac{1}{17} = 0.2017$	$0.0068 + \frac{1}{17^2} = 0.0103$	0.1015
$10 \le t < 13$	$0.2017 + \frac{1}{15} = 0.2683$	$0.0103 + \frac{1}{15^2} = 0.0147$	0.1212
$13 \le t < 16$	$0.2683 + \frac{1}{12} = 0.3517$	$0.0147 + \frac{1}{12^2} = 0.0217$	0.1473
$16 \le t < 22$	$0.3517 + \frac{1}{11} = 0.4426$	$0.0217 + \frac{1}{11^2} = 0.0299$	0.1729
$22 \le t < 23$	$0.4426 + \frac{1}{7} = 0.5854$	$0.0299 + \frac{1}{7^2} = 0.0503$	0.2243
$23 \le t < 35$	$0.5854 + \frac{1}{6} = 0.7521$	$0.0503 + \frac{1}{6^2} = 0.0781$	0.2795

event. Here, one plots the Nelson–Aalen estimator on special paper so that, if a given parametric model fits the data, the resulting graph will be approximately linear. For example, a plot of $\tilde{H}(t)$ versus t will be approximately linear if the exponential distribution, with hazard rate λ, fits the data. The use of the Nelson–Aalen estimators in model identification is discussed further in Chapter 12.

A second use of the Nelson–Aalen estimator is in providing crude estimates of the hazard rate $h(t)$. These estimates are the slope of the Nelson–Aalen estimator. Better estimates of the hazard rate are obtained by smoothing the jump sizes of the Nelson–Aalen estimator with a parametric kernel (see Chapter 6).

EXAMPLE 4.1 *(continued)* The construction of the Nelson–Aalen estimator of the cumulative hazard and its estimated variance for the 6-MP group is given in Table 4.2.

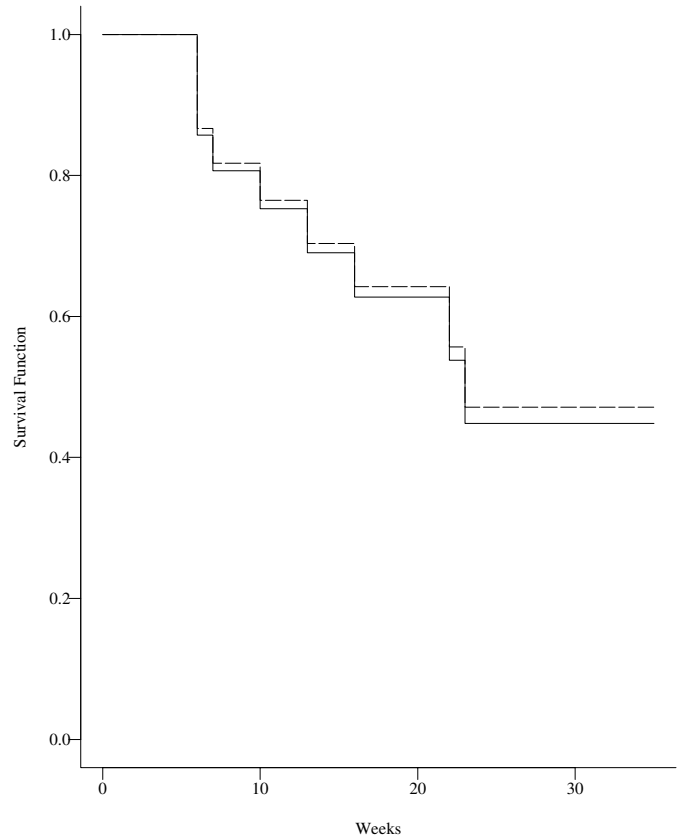

Figure 4.1A *Comparison of the Nelson–Aalen (------) and Product-Limit (————) estimates of the survival function for the 6-MP group.*

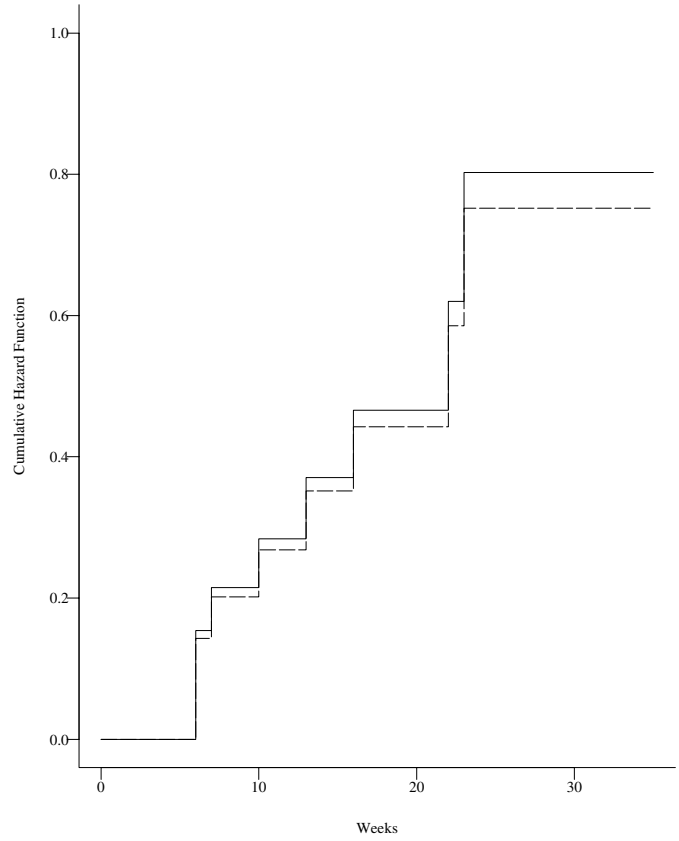

Figure 4.1B *Comparison of the Nelson–Aalen (------) and Product-Limit (————) estimates of the cumulative hazard rate for the 6-MP group.*

Figure 4.1A shows the two estimates of the survival function for the 6-MP data and Figure 4.1B the two estimates of the cumulative hazard rate. Note that all estimates are step functions with jumps at the observed deaths.

EXAMPLE 4.2

To illustrate the use of the Product-Limit estimator and the Nelson–Aalen estimator in providing summary information about survival, consider the data on the efficiency of a bone marrow transplant in acute leukemia. Using the data reported in section 1.3, we shall focus on the disease-free survival probabilities for ALL, AML low risk and AML high risk patients. An individual is said to be disease-free at a given time after transplant if that individual is alive without the recurrence of leukemia. The event indicator for disease-free survival is $\delta_3 = 1$ if the individual has died or has relapsed ($\delta_3 = \max(\delta_1, \delta_2)$ in Table D.1 of Appendix D).

The days on study for a patient is the smaller of their relapse or death time.

Table 4.3 shows the calculations needed for constructing the estimated survival function and hazard rate for the ALL group. Similar calculations for the two AML groups are left as an exercise.

Figure 4.2 shows a plot of the estimated disease-free survival curves (4.2.1) for the three groups. In this figure, first notice that the curves end at different points, because the largest times on study are different for the three groups (2081 days for ALL, 2569 for AML low risk, and 2640 for AML high risk). Secondly, the figure suggests that AML low risk patients have the best and AML high risk patients the least favorable prognosis. The three year disease-free survival probabilities are 0.3531 ($SE = 0.0793$) for the ALL group; 0.5470 ($SE = 0.0691$) for the AML

TABLE 4.3

Estimators of the Survival Function and Cumulative Hazard Rate for ALL Patients

			Product-Limit Estimator		*Nelson–Aalen Estimator*	
t_i	d_i	Y_i	$\hat{S}(t_i)$	$\sqrt{\hat{V}[\hat{S}(t_i)]}$	$\tilde{H}(t_i)$	$\sigma_H(t_i)$
1	1	38	0.9737	0.0260	0.0263	0.0263
55	1	37	0.9474	0.0362	0.0533	0.0377
74	1	36	0.9211	0.0437	0.0811	0.0468
86	1	35	0.8947	0.0498	0.1097	0.0549
104	1	34	0.8684	0.0548	0.1391	0.0623
107	1	33	0.8421	0.0592	0.1694	0.0692
109	1	32	0.8158	0.0629	0.2007	0.0760
110	1	31	0.7895	0.0661	0.2329	0.0825
122	2	30	0.7368	0.0714	0.2996	0.0950
129	1	28	0.7105	0.0736	0.3353	0.1015
172	1	27	0.6842	0.0754	0.3723	0.1081
192	1	26	0.6579	0.0770	0.4108	0.1147
194	1	25	0.6316	0.0783	0.4508	0.1215
230	1	23	0.6041	0.0795	0.4943	0.1290
276	1	22	0.5767	0.0805	0.5397	0.1368
332	1	21	0.5492	0.0812	0.5873	0.1449
383	1	20	0.5217	0.0817	0.6373	0.1532
418	1	19	0.4943	0.0819	0.6900	0.1620
466	1	18	0.4668	0.0818	0.7455	0.1713
487	1	17	0.4394	0.0815	0.8044	0.1811
526	1	16	0.4119	0.0809	0.8669	0.1916
609	1	14	0.3825	0.0803	0.9383	0.2045
662	1	13	0.3531	0.0793	1.0152	0.2185
2081	0	1	0.3531	0.0793	1.0152	0.2185

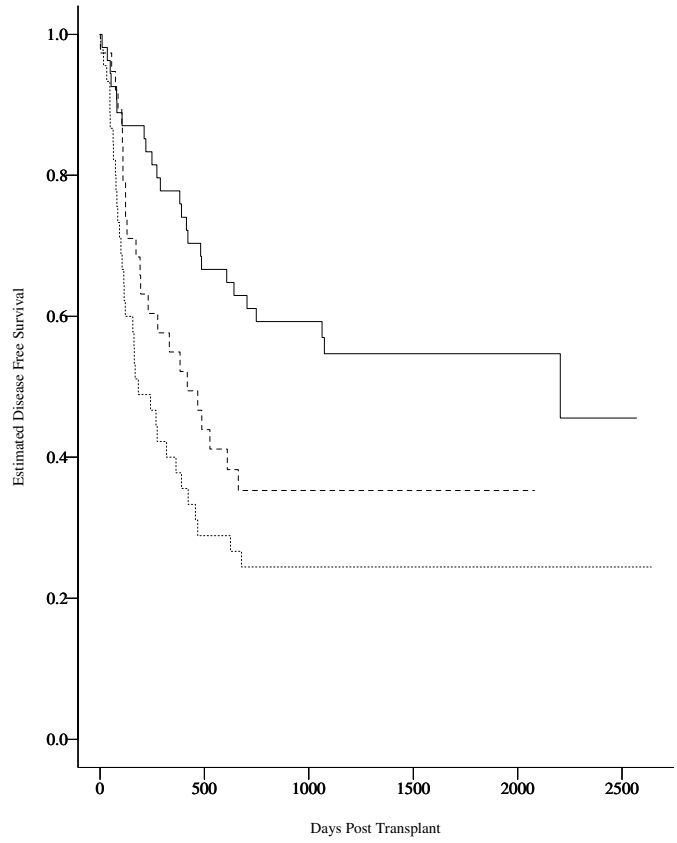

Figure 4.2 *Estimated disease free survival for the 137 bone marrow transplant patients. AML-Low risk (————); AML-High risk (------); ALL (— — —)*

low risk group; and 0.2444 ($SE = 0.0641$) for the AML high risk group. Whether these apparent differences are statistically significant will be addressed in later sections.

Figure 4.3 is a plot of the estimated cumulative hazard rates (4.2.3) for the three disease groups. Again, this plot shows that AML high risk patients have the highest combined relapse and death rate, whereas AML low risk patients have the smallest rate. For each disease group, the cumulative hazard rates appear to be approximately linear in the first two years, suggesting that the hazard rate is approximately constant. A crude estimate of these constant hazard rates is the slopes of the Nelson–Aalen estimators. These estimates give a rate of about 0.04 events per month for ALL patients, 0.02 events per month for AML low risk patients, and 0.06 events per month for AML high risk patients.

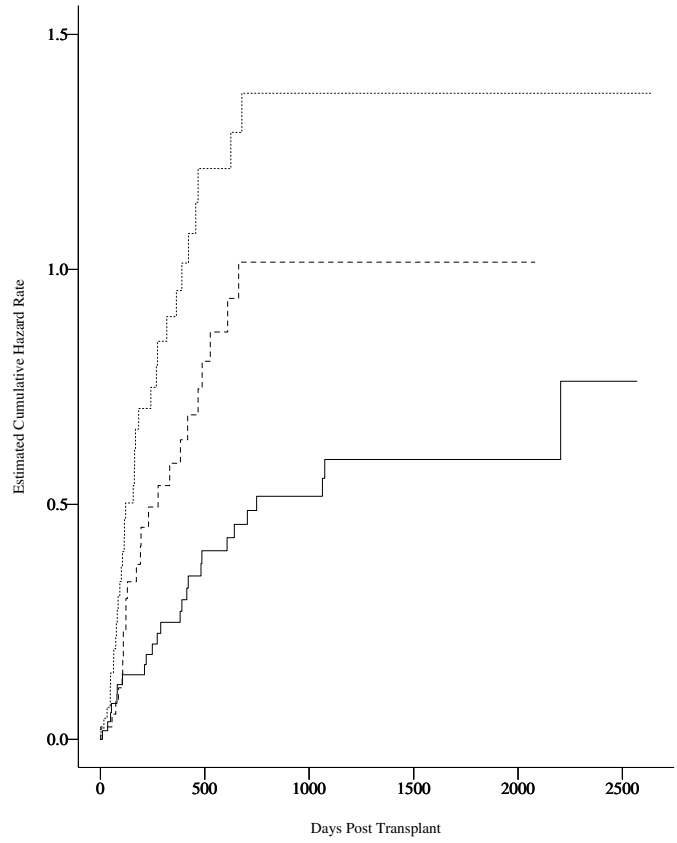

Figure 4.3 *Estimated cumulative hazard rate for the 137 bone marrow transplant patients. AML-Low risk (————); AML-High risk (------); ALL (— — —)*

Practical Notes

1. Both the Nelson–Aalen and Product-Limit estimator are based on an assumption of noninformative censoring which means that knowledge of a censoring time for an individual provides no further information about this person's likelihood of survival at a future time had the individual continued on the study. This assumption would be violated, for example, if patients with poor prognosis were routinely censored. When this assumption is violated, both estimators are estimating the wrong function and the investigator can be appreciably misled. See Klein and Moeschberger (1984) for details.
2. The Kaplan–Meier estimator of the survival function is well defined for all time points less than the largest observed study time $t_{\max}$. If the largest study time corresponds to a death time, then, the estimated

survival curve is zero beyond this point. If the largest time point is censored, the value of $S(t)$ beyond this point is undetermined because we do not know when this last survivor would have died if the survivor had not been censored. Several nonparametric suggestions have been made to account for this ambiguity. Efron (1967) suggests estimating $\hat{S}(t)$ by 0 beyond the largest study time $t_{\max}$. This corresponds to assuming that the survivor with the largest time on study would have died immediately after the survivor's censoring time and leads to an estimator which is negatively biased. Gill (1980) suggests estimating $\hat{S}(t)$ by $\hat{S}(t_{\max})$ for $t > t_{\max}$, which corresponds to assuming this individual would die at ∞ and leads to an estimator which is positively biased. Although both estimators have the same large-sample properties and converge to the true survival function for large samples, a study of small-sample properties of the two estimators by Klein(1991) shows that Gill's version of the Kaplan–Meier is preferred.

3. The two nonparametric techniques for estimation beyond $t_{\max}$ correspond to the two most extreme situations one may encounter. Brown, Hollander, and Kowar (1974) suggest completing the tail by an exponential curve picked to give the same value of $S(t_{\max})$. The estimated survival function for $t > t_{\max}$ is given by $\hat{S}(t) = \exp\{t \ln[\hat{S}(t_{\max})]/t_{\max}\}$. For the data in Example 4.2, this method yields estimates of $\hat{S}(t) = \exp(-0.0005t)$ for $t > 2081$ days for the ALL Group; $\hat{S}(t) = \exp(-0.00035t)$ for $t > 2569$ days for the AML low risk group; and $\hat{S}(t) = \exp(-0.000053t)$ for $t > 2640$ for the AML high risk group. Based on these estimates, Figure 4.4 shows a comparison of the disease-free survival of three-risk groups for the first eight years after transplant. Moeschberger and Klein (1985) have extended these techniques to allow using the more flexible Weibull distribution to complete the tail of the Product-Limit estimator.
4. An alternative estimator of the variance of $\hat{S}(t)$, due to Aalen and Johansen (1978) is given by

$$\tilde{V}[\hat{S}(t)] = \hat{S}(t)^2 \sum_{t_i \le t} \frac{d_i}{Y_i^2}. \tag{4.2.5}$$

Both this estimator and Greenwood's estimator tend to underestimate the true variance of the Kaplan–Meier estimator for small to moderate samples. On average, Greenwood's estimator tends to come closest to the true variance and has a smaller variance except when Y_i is very small (see Klein, 1991).

5. An alternate estimator of the variance of $\tilde{H}(t)$ is found in Klein (1991). This estimator is given as

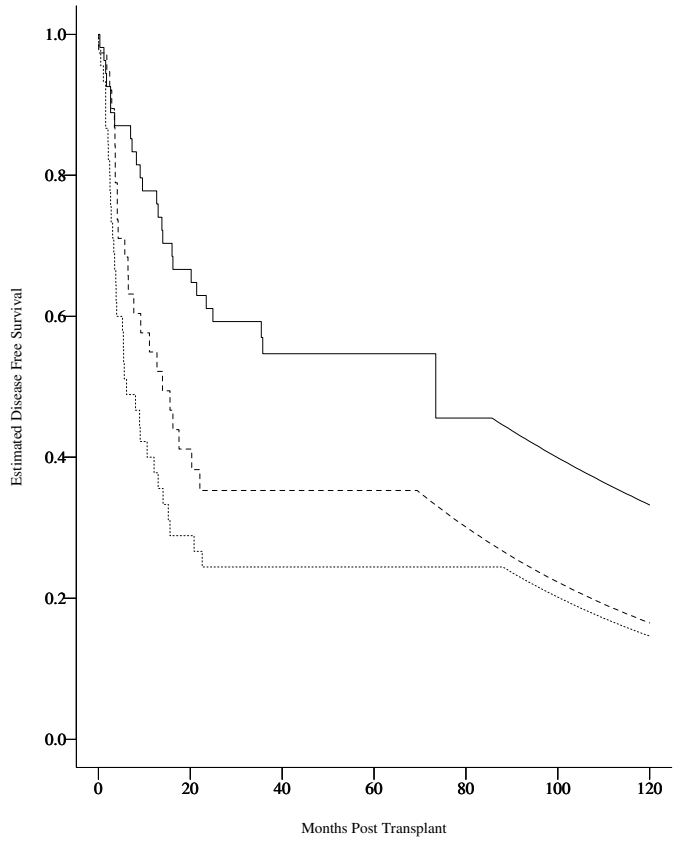

Figure 4.4 *Estimated disease free survival for the 137 bone marrow transplant patients using the Brown-Hollander-Kowar tail estimate. AML-low risk* (———); *AML-high risk* (------); *ALL* (—— ——)

$$\hat{V}[\tilde{H}(t)] = \sum_{t_i \le t} \frac{(Y_i - d_i)d_i}{Y_i^3} \quad . \tag{4.2.6}$$

This estimator tends to be too small, whereas the estimator (4.2.4) tends to be too large. The estimator (4.2.4) has a smaller bias than (4.2.6) and is preferred.

6. An estimator of the variability of the Nelson–Aalen estimator of the survival function $\hat{S}(t)$ is found by substituting $\tilde{S}(t)$ for $\hat{S}(t)$ in either Eq. 4.2.2 or 4.2.5.
7. When there is no censoring, the Product-Limit estimator reduces to the empirical survival function.
8. The statistical packages SAS, BMDP, SPSS, and S-Plus provide procedures for computing the Product-Limit estimator and the estimated

cumulative hazard rate based on this statistic. S-Plus also provides the Nelson–Aalen estimator of the survival function and an estimate of its variance using (4.2.5).

Theoretical Notes

1. The Product-Limit estimator was first constructed by using a *reduced-sample* approach. In this approach, note that, because events are only observed at the times t_i, $S(t)$ should be a step function with jumps only at these times, there being no information on events occurring at other times. We will estimate $S(t)$ by a discrete distribution with mass at the time points $t_1, t_2, \ldots, t_D$. We can estimate the $Pr[T > t_i \mid T \geq t_i]$ as the fraction of individuals who are at risk at time t_i but who do not die at this time, that is

$$\hat{P}r[T > t_i \mid T \geq t_i] = \frac{Y_i - d_i}{Y_i}, \text{for } i = 1, 2, \ldots, D.$$

   To estimate $S(t_i)$, recall that

$$\begin{aligned} S(t_i) &= \frac{S(t_i)}{S(t_{i-1})}\frac{S(t_{i-1})}{S(t_{i-2})} \cdots \frac{S(t_2)}{S(t_1)}\frac{S(t_1)}{S(0)}S(0) \\ &= P[T > t_i \mid T \geq t_i]P[T > t_{i-1} \mid T \geq t_{i-1}] \ldots \\ &\quad P[T > t_2 \mid T \geq t_1]P[T > t_1 \mid T \geq t_1], \end{aligned}$$

   because $S(0) = 1$ and, for a discrete distribution, $S(t_{i-1}) = Pr[T > t_{i-1}] = Pr[T \geq t_i]$. Simplifying this expression yields the Product-Limit estimator (4.2.1).
2. *Redistribute to the right algorithm.* This derivation is best explained by an example. Suppose we have ten individuals on study with the following times on study (+ denotes a censored observation): 3, 4, 5+ ,6, 6+, 8+, 11, 14, 15, 16+. We start the algorithm by assigning mass $1/n$ to each observation (the estimator we would have if there was no censoring). Now, start at the far left and take the mass at each censored observation and redistribute it equally to each observation greater than this value. (Here censored observations tied with events are treated as being just to the right of the event.) This process is repeated until the largest observation is reached. The survival curve obtained from this final set of probabilities is the Kaplan–Meier estimate. If the largest observation is censored, the mass can be either left at that point, so that the Kaplan-Meier estimator drops to zero, or redistributed to $+\infty$, so that the curve is constant beyond this value.
3. *Self Consistency.* If we had no censored observations, the estimator of the survival function at a time t is the proportion of observations

Data Points	*Step 0*	*Step 1*	*Step 2*	*Step 3*	$S(t)$
3	$\frac{1}{10}$	0.100	0.100	0.100	0.900
4	$\frac{1}{10}$	0.100	0.100	0.100	0.800
5+	$\frac{1}{10}$	0.000	0.000	0.000	0.800
6	$\frac{1}{10}$	$\frac{1}{10} + \left(\frac{1}{7}\right)\frac{1}{10} = 0.114$	0.114	0.114	0.686
6+	$\frac{1}{10}$	$\frac{1}{10} + \left(\frac{1}{7}\right)\frac{1}{10} = 0.114$	0.000	0.000	0.686
8+	$\frac{1}{10}$	$\frac{1}{10} + \left(\frac{1}{7}\right)\frac{1}{10} = 0.114$	$0.114 + \frac{1}{5}0.114 = 0.137$	0.000	0.686
11	$\frac{1}{10}$	$\frac{1}{10} + \left(\frac{1}{7}\right)\frac{1}{10} = 0.114$	$0.114 + \frac{1}{5}0.114 = 0.137$	$0.137 + \frac{1}{4}0.137 = 0.171$	0.515
14	$\frac{1}{10}$	$\frac{1}{10} + \left(\frac{1}{7}\right)\frac{1}{10} = 0.114$	$0.114 + \frac{1}{5}0.114 = 0.137$	$0.137 + \frac{1}{4}0.137 = 0.171$	0.343
15	$\frac{1}{10}$	$\frac{1}{10} + \left(\frac{1}{7}\right)\frac{1}{10} = 0.114$	$0.114 + \frac{1}{5}0.114 = 0.137$	$0.137 + \frac{1}{4}0.137 = 0.171$	0.171
16+	$\frac{1}{10}$	$\frac{1}{10} + \left(\frac{1}{7}\right)\frac{1}{10} = 0.114$	$0.114 + \frac{1}{5}0.114 = 0.137$	$0.137 + \frac{1}{4}0.137 = 0.171$	0.000*

*Efron's Estimator

which are larger than t, that is,

$$\hat{S}(t) = \frac{1}{n}\sum_{i=1}^{n} \phi(X_i)$$

where

$$\phi(y) = 1 \text{ if } y > t; 0, \text{ if } y \leq t.$$

For right-censored data, we want to construct our estimator in a similar manner by redefining the scoring function ϕ. Let $T_1, T_2, \ldots, T_n$ be the observed times on study. If T_i is a death time, we know with certainty whether T_i is smaller or greater than t. If T_i is a censored time greater than or equal to t, then, we know that the true death time must be larger than t because it is larger than T_i for this individual. For a censored observation less than t, we do not know if the corresponding death time is greater than t because it could fall between T_i and t. If we knew $S(t)$, we could estimate the probability of this censored observation being larger than t by $S(t)/S(T_i)$. Using these revised scores, we will call an estimator $\hat{S}(t)$ a self-consistent

estimator of S if

$$\hat{S}(t) = \frac{1}{n}\left[\sum_{T_i > t} \phi(T_i) + \sum_{\delta_i = 0, T_i \le t} \frac{\hat{S}(t)}{\hat{S}(T_i)}\right]. \tag{4.2.7}$$

To find $\hat{S}(t)$ from this formula one starts with any estimate of S and substitutes this in the right hand side of (4.2.7) to get an updated estimate of S. This new estimate of $\hat{S}(t)$ is, then, used in the next step to obtain a revised estimate. This procedure continues until convergence. Efron (1967) shows that the final estimate of S is exactly the Product-Limit estimator for t less than the largest observed time.

4. Both the Product-Limit estimator and the Nelson–Aalen estimator can be derived using the theory of counting processes. Details of this construction can be found in Andersen et al. (1993) or Fleming and Harrington (1991). An outline of this approach is found in section 3.6.
5. Under certain regularity conditions, one can show that the Nelson–Aalen estimator and the Product-Limit estimator are nonparametric maximum likelihood estimators.
6. Both the Product-Limit and Nelson–Aalen estimators of either the survival function or the cumulative hazard rate are consistent. The statistics are asymptotically equivalent.
7. The Nelson–Aalen estimator of the cumulative hazard rate is the first term in a Taylor series expansion of minus the logarithm of the Product-Limit estimator.
8. Small-sample-size properties of the Product-Limit estimator have been studied by Guerts (1987) and Wellner (1985). Small sample size properties of the variance estimators for the Product-Limit estimator and the Nelson–Aalen estimator are found in Klein (1991).
9. Under suitable regularity conditions, both the Nelson–Aalen and Product-Limit estimators converge weakly to Gaussian processes. This fact means that for fixed t, the estimators have an approximate normal distribution.

4.3 Pointwise Confidence Intervals for the Survival Function

The Product-Limit estimator provides a summary estimate of the mortality experience of a given population. The corresponding standard error provides some limited information about the precision of the estimate. In this section, we shall use these estimators to provide confidence intervals for the survival function at a fixed time t_o. The intervals are

constructed to assure, with a given confidence level $1 - \alpha$ that the true value of the survival function, at a predetermined time t_o, falls in the interval we shall construct.

Before introducing the confidence intervals, we need some additional notation. Let $\sigma_S^2(t) = \hat{V}[\hat{S}(t)]/\hat{S}^2(t)$. Note that $\sigma_S^2(t)$ is the sum in Greenwood's formula (4.2.2).

The most commonly used $100 \times (1 - \alpha)\%$ confidence interval for the survival function at time t_o, termed the linear confidence interval, is defined by

$$\hat{S}(t_o) - Z_{1-\alpha/2}\sigma_S(t_o)\hat{S}(t_o), \hat{S}(t_o) + Z_{1-\alpha/2}\sigma_S(t_o)\hat{S}(t_o) \tag{4.3.1}$$

Here $Z_{1-\alpha/2}$ is the $1 - \alpha/2$ percentile of a standard normal distribution. This is the confidence interval routinely constructed by most statistical packages.

Better confidence intervals can be constructed by first transforming $\hat{S}(t_o)$. These improved estimators were proposed by Borgan and Liestøl (1990). The first transformation suggested is a log transformation (see Theoretical Note 4) of the cumulative hazard rate. The $100 \times (1 - \alpha)\%$ log-transformed confidence interval for the survival function at t_o is given by

$$[\hat{S}(t_o)^{1/\theta}, \hat{S}(t_o)^{\theta}], \text{where } \theta = \exp\left\{\frac{Z_{1-\alpha/2}\sigma_S(t_o)}{\ln[\hat{S}(t_o)]}\right\}. \tag{4.3.2}$$

Note that this interval is not symmetric about the estimate of the survival function.

The second transformation is an arcsine-square root transformation of the survival function which yields the following $100 \times (1 - \alpha)\%$ confidence interval for the survival function:

$$\sin^2\left\{\max\left[0, \arcsin(\hat{S}(t_o)^{1/2}) - 0.5Z_{1-\alpha/2}\sigma_S(t_o)\left(\frac{\hat{S}(t_o)}{1-\hat{S}(t_o)}\right)^{1/2}\right]\right\}$$
$$\leq S(t_o) \leq \tag{4.3.3}$$
$$\sin^2\left\{\min\left[\frac{\pi}{2}, \arcsin(\hat{S}(t_o)^{1/2}) + 0.5Z_{1-\alpha/2}\sigma_S(t_o)\left(\frac{\hat{S}(t_o)}{1-\hat{S}(t_o)}\right)^{1/2}\right]\right\}.$$

EXAMPLE 4.2 *(continued)* To illustrate these confidence intervals, we shall use the estimated disease-free survival function and cumulative hazard rate for ALL patients in Table 4.3. Note that at 1 year (365 days) the estimated survival function S(365) was found to be 0.5492 with an estimated variance of 0.0812^2. Thus, $\sigma_S^2(365) = (0.0812/0.5492)^2 = 0.1479^2$. A 95% linear confidence interval for the survival function at year one is $0.5492 \pm 1.96 \times 0.1479 \times 0.5492 = (0.3900, 0.7084)$.

To find the 95% log-transformed confidence interval for the one year survival function, we find that $\theta = \exp[\frac{1.96 \times 0.1479}{\ln(0.5492)}] = 0.6165$, so that the interval is $(0.54921^{1/0.6165}, 0.5492^{0.6165}) = (0.3783, 0.6911)$.

The 95% arcsine-square root transformation confidence interval for the one year survival function is

$$\sin^2\left\{\max\left[0, \arcsin(0.5492^{1/2}) - 0.5 \times 1.96 \times 0.1479 \times \left(\frac{0.5492}{1-0.5492}\right)^{1/2}\right]\right\}$$

to

$$\sin^2\left\{\min\left[\frac{\pi}{2}, \arcsin(0.5492^{1/2}) + 0.5 \times 1.96 \times 0.1479 \times \left(\frac{0.5492}{1-0.5492}\right)^{1/2}\right]\right\}$$

$= (0.3903, 0.7032)$.

Table 4.4 shows the three possible 95% confidence intervals that can be constructed for the disease-free survival function for each of the three risk groups presented in Figures 4.2. We can see that AML high risk patients have a smaller chance of surviving beyond one year than the AML low risk patients.

TABLE 4.4
95% Confidence Intervals for Disease-Free Survival One Year After Transplant

	ALL	*AML low risk*	*AML high risk*
$\hat{S}(365)$	0.5492	0.7778	0.3778
$\hat{V}[\hat{S}(365)]$	0.0812^2	0.0566^2	0.0723^2
$\sigma_s(365)$	0.1479	0.0728	0.1914
Linear confidence interval for $S(365)$	0.3900, 0.7084	0.6669, 0.8887	0.2361, 0.5195
Log-transformed confidence interval for $S(365)$	0.3783, 0.6911	0.6419, 0.8672	0.2391, 0.5158
Arcsine square-root confidence interval for $S(365)$	0.3903, 0.7032	0.6583, 0.8776	0.2433, 0.5227

Practical Notes

1. Bie et al. (1987) have presented $100(1-\alpha)\%$ pointwise confidence intervals for the cumulative hazard function. Similar to the confi-

dence intervals constructed for the survival function, there are three possible intervals, which correspond to the three transformations of the cumulative hazard function. The intervals are

Linear:

$$\tilde{H}(t_o) - Z_{1-\alpha/2}\sigma_H(t_o),\ \tilde{H}(t_o) + Z_{1-\alpha/2}\sigma_H(t_o). \tag{4.3.4}$$

Log-Transformed

$$[\tilde{H}(t_o)/\phi, \phi\tilde{H}(t_o)] \text{ where } \phi = \exp[\frac{Z_{1-\alpha/2}\sigma_H(t_o)}{\tilde{H}(t_o)}]. \tag{4.3.5}$$

Arcsine-Square Root Transformed

$$\begin{aligned} &-2\ln\Big\{\sin\Big[\min\Big(\frac{\pi}{2}, \arcsin[\exp\{-\tilde{H}(t_o)/2\}] \\ &\quad +0.5Z_{1-\alpha/2}\sigma_H(t_o)\{\exp\{\tilde{H}(t_o)\} - 1\}^{-1/2}\Big)\Big]\Big\} \\ &\le H(t_o) \le \\ &-2\ln\{\sin[\max(0, \arcsin[\exp\{-\tilde{H}(t_o)/2\}] \\ &\quad -0.5Z_{1-\alpha/2}\sigma_H(t_o)\{\exp\tilde{H}(t_o)\} - 1\}^{-1/2})]\} \end{aligned} \tag{4.3.6}$$

Using the data in Example 4.2, we have the following 95% confidence intervals for the cumulative hazard rate at one year after transplant:

	ALL	*AML low risk*	*AML high risk*
Linear confidence interval for H(365)	0.3034, 0.8713	0.1076, 0.3898	0.5875, 1.3221
Log-transformed confidence interval for H(365)	0.3622, 0.9524	0.1410, 0.4385	0.6499, 1.4028
Arcsin square root confidence interval for H(365)	0.3451, 0.9217	0.1293, 0.4136	0.6366, 1.3850

2. Borgan and Liestøl (1990) have shown that both the log-transformed and arcsine-square root transformed confidence intervals for S perform better than the usual linear confidence interval. Both give about the correct coverage probability for a 95% interval for samples as small as 25 with as much as 50% censoring except in the extreme right-hand tail where there will be little data. The sample size needed for the standard linear confidence interval to have the correct coverage probability is much larger. For very small samples, the arcsine-square root interval tends to be a bit conservative in that the actual

coverage probability is a bit greater than $(1 - \alpha)$, whereas, for the log-transformed interval, the coverage probability is a bit smaller than $(1 - \alpha)$. The coverage probability for the linear interval in these cases is much smaller than $(1 - \alpha)$. Similar observations were made by Bie et al. (1987) for the corresponding interval estimates of the cumulative hazard rate. For very large samples, the three methods are equivalent.

3. Alternative confidence intervals for the cumulative hazard rate can be found by taking (minus) the natural logarithm of the confidence intervals constructed for the survival function. Similarly the exponential of (minus) the confidence limits for the cumulative hazard yields a confidence interval for the survival function.
4. Both the log-transformed and arcsine-square root transformed confidence intervals, unlike the linear interval, are not symmetric about the point estimator of the survival function or cumulative hazard rate. This is appropriate for small samples where the point estimators are biased and the distribution of the estimators is skewed.
5. The confidence intervals constructed in this section are valid only at a single point t_o. A common incorrect use of these intervals is to plot them for all values of t and interpret the curves obtained as a confidence band, that is, these curves are interpreted as having, for example, 95% confidence that the *entire* survival function lies within the band. The bands obtained this way are too narrow to make this inference. The proper bands are discussed in the following section.
6. Confidence intervals for the survival function are available in the S-Plus routine surv.fit. The intervals can be constructed using either the linear or the log-transformed method.

Theoretical Notes

1. Construction of the linear confidence intervals follows directly from the asymptotic normality of the Product-Limit or Nelson–Aalen estimators.
2. The log-transformed interval was first suggested by Kalbfleisch and Prentice (1980).
3. The arcsine-square root transformed interval was first suggested by Nair (1984).
4. The "log"-transformed confidence interval is based on first finding a confidence interval for the log of the cumulative hazard function. This is sometimes called a log-log transformed interval since the cumulative hazard function is the negative log of the survival function.

4.4 Confidence Bands for the Survival Function

In section 4.3, pointwise confidence intervals for the survival function were presented. These intervals are valid for a single fixed time at which the inference is to be made. In some applications it is of interest to find upper and lower confidence bands which guarantee, with a given confidence level, that the survival function falls within the band for all t in some interval, that is, we wish to find two random functions $L(t)$ and $U(t)$, so that $1-\alpha = Pr[L(t) \leq S(t) \leq U(t)$, for all $t_L \leq t \leq t_U]$. We call such a $[L(t), U(t)]$ a $(1-\alpha) \times 100\%$ confidence band for $S(t)$.

We shall present two approaches to constructing confidence bands for $S(t)$. The first approach, originally suggested by Nair (1984), provides confidence bounds which are proportional to the pointwise confidence intervals discussed in section 4.3. These bands are called the equal probability or *EP* bands. To implement these bands we pick $t_L < t_U$ so that t_L is greater than or equal to the smallest observed event time and t_U is less than or equal to the largest observed event time. To construct confidence bands for $S(t)$, based on a sample of size n, define

$$a_L = \frac{n\sigma_S^2(t_L)}{1 + n\sigma_S^2(t_L)} \tag{4.4.1}$$

and

$$a_U = \frac{n\sigma_S^2(t_U)}{1 + n\sigma_S^2(t_U)} .$$

The construction of the EP confidence bands requires that $0 < a_L < a_U < 1$.

To construct a $100(1-\alpha)\%$ confidence band for $S(t)$ over the range $[t_L, t_U]$, we, first, find a confidence coefficient, $c_\alpha(a_L, a_U)$ from Table C.3 in Appendix C. As in the case of $100(1-\alpha)\%$ pointwise confidence intervals at a fixed time, there are three possible forms for the confidence bands. The three bands are the linear bands, the log-transformed bands, and the arcsine-square root transformed bands expressed as follows:

Linear:

$$\hat{S}(t) - c_\alpha(a_L, a_U)\sigma_S(t)\hat{S}(t), \hat{S}(t) + c_\alpha(a_L, a_U)\sigma_S(t)\hat{S}(t). \tag{4.4.2}$$

Log-Transformed:

$$(\hat{S}(t)^{1/\theta}, \hat{S}(t)^{\theta}),$$

$$\text{where } \theta = \exp\left[\frac{c_\alpha(a_L, a_U)\sigma_S(t)}{\ln[\hat{S}(t)]}\right] . \tag{4.4.3}$$

Arcsine-Square Root Transformed:

$$\sin^2\{\max[0, \arcsin\{\hat{S}(t)^{1/2}\} - 0.5c_\alpha(a_L, a_U)\sigma_S(t)[\hat{S}(t)/(1-\hat{S}(t))]^{1/2}]\}$$
$$\leq S(t) \leq \tag{4.4.4}$$
$$\sin^2\left\{\min\left[\frac{\pi}{2}, \arcsin\{\hat{S}(t)^{1/2}\} + 0.5c_\alpha(a_L, a_U)\sigma_S(t)[\hat{S}(t)/(1-\hat{S}(t))]^{1/2}\right]\right\}.$$

EXAMPLE 4.2

(continued) To illustrate these confidence intervals, we shall use the estimated disease-free survival function for ALL patients in Table 4.3. We construct confidence bands for $S(t)$ over the range $100 \leq t \leq 600$ days. Here, we have $\sigma_S^2(100) = \sigma_S^2(86) = 0.0498^2/0.8947^2 = 0.0031$ and $\sigma_S^2(600) = \sigma_S^2(526) = 0.0809^2/0.4119^2 = 0.0386$. From 4.4.1 we find $a_L = 38(0.0031)/[1 + 38(0.0031)] = 0.1$ and $a_U = 38(0.0386)/[1 + 38(0.0386)] = 0.6$. For a 95% confidence band, we find, from Table C.3 in Appendix C, that $c_{05}(0.1, 0.6) = 2.8826$.

Table 4.5 shows the three 95% confidence bands for the survival function based on the EP method. Note that the calculation of the entries in this table is identical to the calculations performed in section 4.3 for the 95% pointwise confidence intervals at day 365 with the exception that the Z coefficient, 1.96 is replaced by the appropriate value from Table C.3 of Appendix C.

An alternate set of confidence bounds has been suggested by Hall and Wellner (1980). These bands are not proportional to the pointwise confidence bounds. For these bounds, a lower limit, t_L, of zero is allowed. To construct a $100 \times (1-\alpha)\%$ confidence band for $S(t)$ over the region $[t_L, t_U]$, we find the appropriate confidence coefficient $k_\alpha(a_L, a_U)$, from Table C.4 of Appendix C. Again, there are three possible forms for the confidence bands. These are the linear bands, the log-transformed bands and the arcsine-transformed bands. These $100 \times (1-\alpha)\%$ confidence bands are expressed as

Linear:

$$\hat{S}(t) - \frac{k_\alpha(a_L, a_U)[1 + n\sigma_S^2(t)]}{n^{1/2}}\hat{S}(t), \quad \hat{S}(t) + \frac{k_\alpha(a_L, a_U)[1 + n\sigma_S^2(t)]}{n^{1/2}}\hat{S}(t). \tag{4.4.5}$$

Log-Transformed:

$$[\hat{S}(t)^{1/\theta}, \hat{S}(t)^\theta], \tag{4.4.6}$$
$$\text{where } \theta = \exp\left\{\frac{k_\alpha(a_L, a_U)[1 + n\sigma_S^2(t)]}{n^{1/2}\ln[\hat{S}(t)]}\right\}.$$

TABLE 4.5
95% EP Confidence Bands for the Disease Free Survival Function

t_i	$\hat{S}(t_i)$	$\sqrt{\hat{V}[\hat{S}(t_i)]}$	σ_S^2	*Linear*		*Log-Transformed*		*Arcsine-Transformed*	
100	0.8947	0.0498	0.0031	0.7511	1.0000	0.6246	0.9740	0.7139	0.9907
104	0.8684	0.0548	0.0040	0.7104	1.0000	0.5992	0.9619	0.6766	0.9812
107	0.8421	0.0592	0.0049	0.6715	1.0000	0.5719	0.9485	0.6408	0.9698
109	0.8158	0.0629	0.0059	0.6345	0.9971	0.5452	0.9339	0.6071	0.9567
110	0.7895	0.0661	0.0070	0.5990	0.9800	0.5188	0.9184	0.5748	0.9421
122	0.7368	0.0714	0.0094	0.5310	0.9426	0.4666	0.8848	0.5130	0.9098
129	0.7105	0.0736	0.0107	0.4983	0.9227	0.4410	0.8670	0.4834	0.8924
172	0.6842	0.0754	0.0121	0.4669	0.9015	0.4162	0.8485	0.4549	0.8739
192	0.6579	0.0770	0.0137	0.4359	0.8799	0.3917	0.8294	0.4270	0.8549
194	0.6316	0.0783	0.0154	0.4059	0.8573	0.3678	0.8097	0.3999	0.8350
230	0.6041	0.0795	0.0173	0.3749	0.8333	0.3431	0.7886	0.3720	0.8137
276	0.5767	0.0805	0.0195	0.3447	0.8087	0.3187	0.7672	0.3448	0.7920
332	0.5492	0.0812	0.0219	0.3151	0.7833	0.2951	0.7451	0.3183	0.7694
383	0.5217	0.0817	0.0245	0.2862	0.7572	0.2719	0.7224	0.2925	0.7462
418	0.4943	0.0819	0.0275	0.2582	0.7304	0.2496	0.6993	0.2675	0.7223
468	0.4668	0.0818	0.0307	0.2310	0.7026	0.2280	0.6753	0.2433	0.6976
487	0.4394	0.0815	0.0344	0.2045	0.6743	0.2069	0.6510	0.2198	0.6723
526	0.4119	0.0809	0.0386	0.1787	0.6451	0.1865	0.6259	0.1970	0.6462
600	0.4119	0.0809	0.0386	0.1787	0.6451	0.1865	0.6259	0.1970	0.6462

TABLE 4.6
95% Hall–Wellner Confidence Bands for the Disease-Free Survival Function

t_i	$\hat{S}(t_i)$	σ_S^2	*Linear*		*Log-Transformed*		*Arcsine-square root Transformed*	
100	0.8947	0.0031	0.6804	1.0000	0.3837	0.9872	0.6050	1.0000
104	0.8684	0.0040	0.6541	1.0000	0.4445	0.9757	0.5966	0.9971
107	0.8421	0.0049	0.6277	1.0000	0.4696	0.9617	0.5824	0.9869
109	0.8158	0.0059	0.6015	1.0000	0.4771	0.9455	0.5652	0.9723
110	0.7895	0.0070	0.5752	1.0000	0.4747	0.9278	0.5459	0.9550
122	0.7368	0.0094	0.5225	0.9511	0.4532	0.8888	0.5034	0.9152
129	0.7105	0.0107	0.4961	0.9249	0.4377	0.8682	0.4810	0.8939
172	0.6842	0.0121	0.4699	0.8985	0.4205	0.8468	0.4582	0.8718
192	0.6579	0.0137	0.4435	0.8723	0.4018	0.8251	0.4349	0.8492
194	0.6316	0.0154	0.4172	0.8460	0.3822	0.8029	0.4114	0.8262
230	0.6041	0.0173	0.3894	0.8188	0.3606	0.7796	0.3864	0.8021
276	0.5767	0.0195	0.3616	0.7918	0.3383	0.7561	0.3612	0.7779
332	0.5492	0.0219	0.3337	0.7647	0.3156	0.7324	0.3359	0.7535
383	0.5217	0.0245	0.3057	0.7377	0.2925	0.7087	0.3104	0.7290
418	0.4943	0.0275	0.2779	0.7107	0.2694	0.6849	0.2851	0.7046
468	0.4668	0.0307	0.2500	0.6836	0.2462	0.6609	0.2599	0.6799
487	0.4394	0.0344	0.2221	0.6567	0.2230	0.6372	0.2347	0.6555
526	0.4119	0.0386	0.1942	0.6296	0.2000	0.6133	0.2097	0.6310
600	0.4119	0.0386	0.1942	0.6296	0.2000	0.6133	0.2097	0.6310

Arcsine-Square Root Transformation:

$$\sin^2\{\max[0, \arcsin\{\hat{S}(t)^{1/2}\} - 0.5\frac{k_\alpha(a_L, a_U)[1 + n\sigma_S^2(t)]}{n^{1/2}}[(\hat{S}(t)/(1 - \hat{S}(t))]^{1/2}]\}$$
$$\leq S(t) \leq \tag{4.4.7}$$
$$\sin^2\left\{\min\left[\frac{\pi}{2}, \arcsin\{\hat{S}(t)^{1/2}\} + 0.5\frac{k_\alpha(a_L, a_U)\{1 + n\sigma_S^2(t)\}}{n^{1/2}}[\hat{S}(t)/(1 - \hat{S}(t))]^{1/2}\right]\right\}.$$

EXAMPLE 4.2 *(continued)* To illustrate the Hall-Wellner confidence bands, again, we consider the disease-free survival estimates for $S(t)$ obtained from the 38 ALL patients in Table 4.3. As for the EP bands, we construct 95%

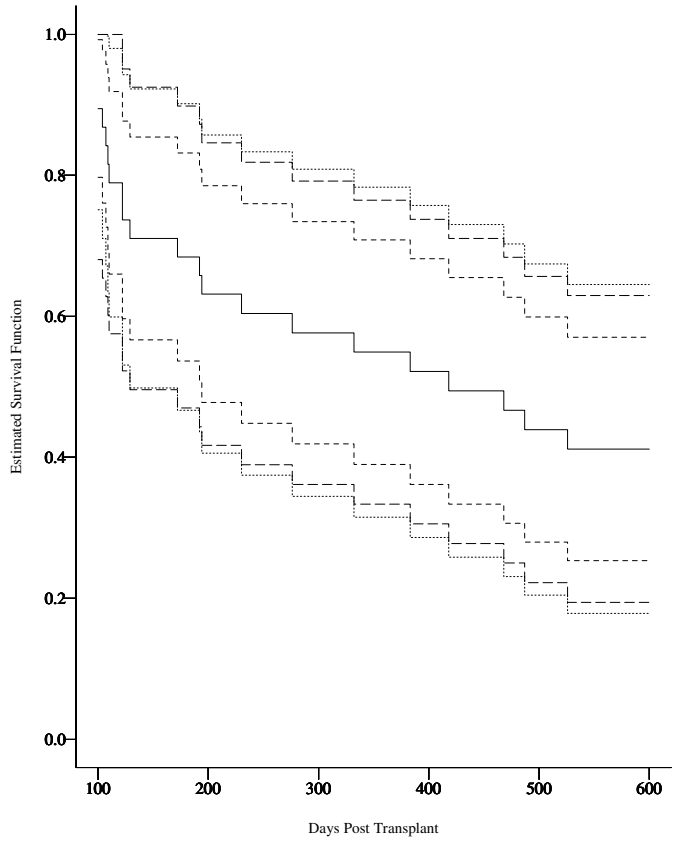

Figure 4.5 *Comparison of 95% pointwise confidence interval, EP confidence band and Hall-Wellner confidence band for the disease free survival function based on the untransformed survival functions for ALL patients. Estimated Survival (————); Pointwise confidence interval (— — —); EP confidence band (------); Hall-Wellner band (——— ———)*

confidence bands for $S(t)$ in the range $100 \leq t \leq 600$. The required confidence coefficient, from Table C.4 of Appendix C, is $k_{05}(0.1, 0.6) = 1.3211$. Table 4.6 shows the Hall-Wellner 95% confidence bands based on the three transformations.

Figures 4.5–4.7 show the 95% confidence bands for the disease-free survival function based on either the EP or Hall–Wellner bands for the three transformations. Also included are the 95% pointwise confidence intervals obtained from the results of section 4.3. These figures show that the Hall–Wellner bands are wider for small t and shorter for large t. Both bands are wider than the curves one obtains by using pointwise confidence intervals.

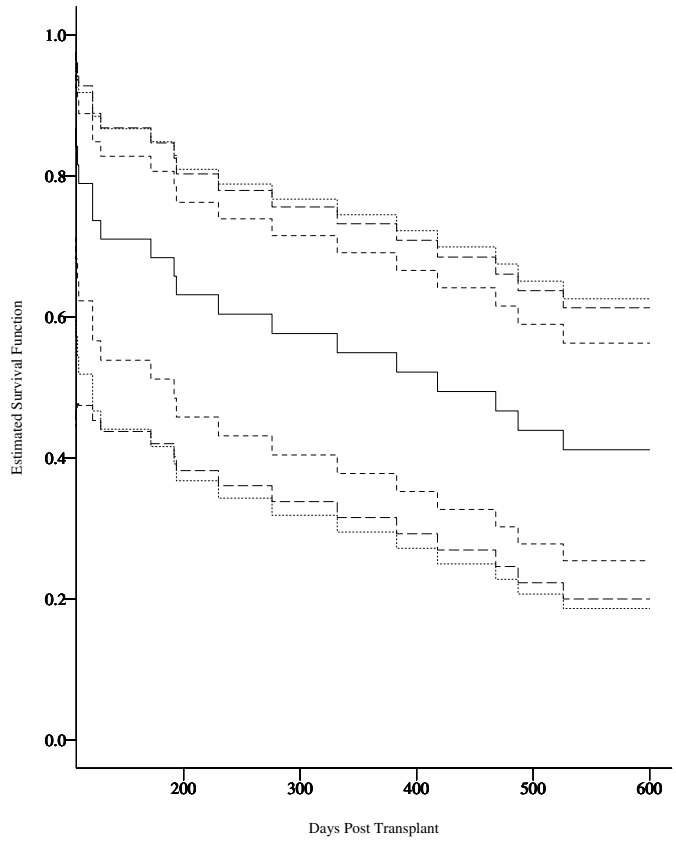

Figure 4.6 *Comparison of 95% pointwise confidence interval, EP confidence band and Hall-Wellner confidence band for the disease free survival function found using the log transformation for ALL patients. Estimated Survival* (————)*; Pointwise confidence interval* (— — —)*; EP confidence band* (------)*; Hall-Wellner band* (——— ———)

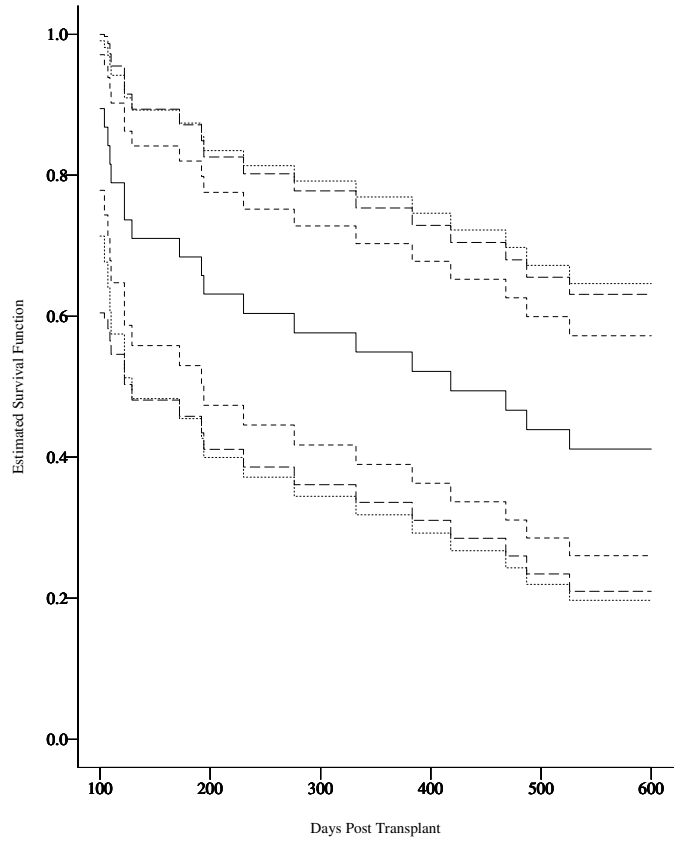

Figure 4.7 *Comparison of 95% pointwise confidence interval, EP confidence band and Hall-Wellner confidence band for the disease free survival function found using the arc sine transformation for ALL patients. Estimated Survival* (————); *Pointwise confidence interval* (— — —); *EP confidence band* (------); *Hall-Wellner band* (— — — —)

Figure 4.8 shows the 95% EP arcsine-square root transformed confidence bands for the three disease categories over the range of 100 to 600 days.

Practical Notes

1. Confidence bands for the cumulative hazard rate can also be constructed by either the EP or Hall–Wellner method. To construct these

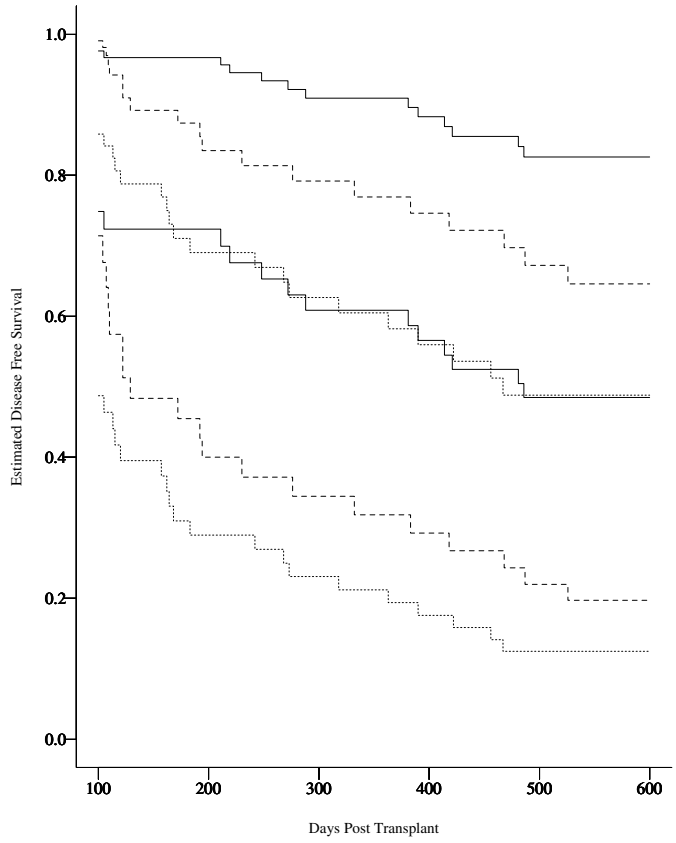

Figure 4.8 *EP confidence bands for the disease free survival function based on the arc sine transformation for bone marrow transplant patients. AML-Low risk (———); AML-High risk (------); ALL (— — —)*

bands, we first compute

$$a_{\mathrm{L}} = \frac{n\sigma_H^2(t_{\mathrm{L}})}{1 + n\sigma_H^2(t_{\mathrm{L}})} \tag{4.4.8}$$

and

$$a_{\mathrm{U}} = \frac{n\sigma_H^2(t_{\mathrm{U}})}{1 + n\sigma_H^2(t_{\mathrm{U}})}.$$

The EP confidence bands, which are valid over the range $t_{\mathrm{L}} \leq t \leq t_{\mathrm{U}}$, with $0 < a_{\mathrm{L}} < a_{\mathrm{U}} < 1$, are found by substituting for $Z_{1-\alpha/2}$ in (4.3.4)–(4.3.6) the appropriate confidence coefficient $c_\alpha(a_{\mathrm{L}}, a_{\mathrm{U}})$ from Table C.3 of Appendix C.

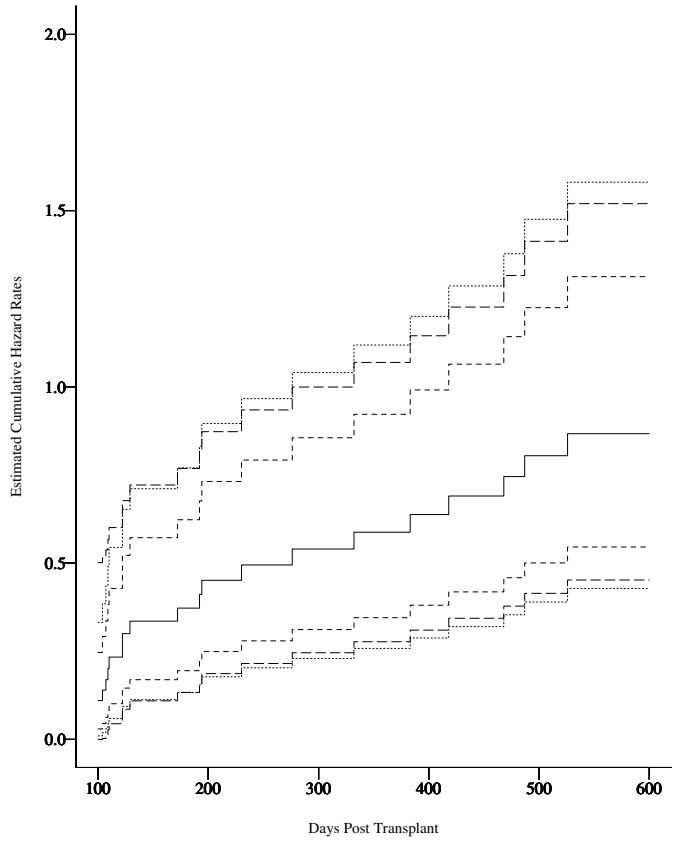

Figure 4.9 *Comparison of 95% pointwise confidence interval, EP confidence band and Hall-Wellner confidence band for the cumulative hazard function found using the arc sine transformation for ALL patients. Estimated Survival* (————)*; Pointwise confidence interval* (— — —)*; EP confidence band* (------)*; Hall-Wellner band* (——— ———)

The Hall–Wellner confidence bands for the cumulative hazard rate are found by substituting $\frac{k_\alpha(a_L, a_U)[1+n\sigma_H^2(t)]}{n^{1/2}}$ for $Z_{1-\alpha/2}\sigma_H(t)$ in (4.3.4)–(4.3.6).

Figure 4.9 shows the 95% arcsine-square root transformed EP and Hall–Wellner confidence bands and the 95% pointwise confidence interval for the cumulative hazard rate of the ALL patients over the interval 100 to 600 days.

2. For the EP bounds for the survival function, Borgan and Liestøl (1990) have shown that the linear confidence band given by formula (4.4.2) performs very poorly when the sample size is small (< 200). The coverage probability for this bound is considerably smaller than the target level. For both the log- and arcsine-square root transformed

bands, the coverage probability is approximately correct for smaller sample sizes. Both seem to give reasonable results for samples with as few as 20 observed events. The arcsine-square root transformed band seems to perform a bit better than the log-transformed interval and is recommended. Similar properties hold for the confidence bands for the cumulative hazard rate as discussed in Bie et al. (1987).

3. For the Hall–Wellner bounds, Borgan and Liestøl (1990) show that all three bands for $S(t)$ perform reasonably well for samples with as few as 20 observed events. For $H(t)$, Bie et al. (1987) show that the performance of the linear bands is poor for small samples, whereas the two transformed bands perform well for relatively small samples.
4. For the confidence bands for $H(t)$, linear EP confidence bands tend to have an upper band which is a bit too low, whereas the log-transformed lower band is too high for small t and the upper band too low for large t. For the EP arcsine-square root band, the majority of the errors occur when the upper boundary is too low. For the HW bands, the majority of the errors occur in the midrange of t.

Theoretical Notes

1. These bounds are based on weak convergence of the Product-Limit estimator or the Nelson–Aalen estimator to a mean zero Gaussian process. The EP bounds are based on the transformation $q(x) = [x(1-x)]^{-1/2}$ of the standardized estimator, whereas for the Hall–Wellner bounds, no transformation of this process is made.
2. The critical values found in Table C.3 of Appendix C are the upper αth fractile of the random variable $U = \sup\{| W^{\circ}(x)[x(1-x)]^{-1/2} |, a_L \leq x \leq a_U\}$, where W° is a standard Brownian bridge (see Nair, 1984). Miller and Siegmund (1982) show that, for large d, $Pr[U \geq d) \cong 4\phi(d)/d + \phi(d)(d - d^{-1}) \log[\frac{a_U(1-a_L)}{a_L(1-a_U)}]$, where $\phi(\,)$ is the standard normal density function.
3. The critical values for the Hall–Wellner bands (Table C.4 of Appendix C) are the upper αth fractile of a Brownian bridge, computed from results in Chung (1986).

4.5 Point and Interval Estimates of the Mean and Median Survival Time

The Product-Limit estimator provides an estimator of the survival function $S(t)$. In section 2.4, we saw that other summary measures of an individual's survival experience, such as the mean or median time to

the event, are functions of the survival function. Nonparametric estimates of these quantities can be obtained in a straightforward manner by substituting the Product-Limit estimator for the unknown survival function in the appropriate formula.

In section 2.4, it was shown that the mean time to the event μ is given by $\mu = \int_0^\infty S(t)dt$. A natural estimator of μ is obtained by substituting $\hat{S}(t)$ for $S(t)$ in this expression. This estimator is appropriate only when the largest observation corresponds to a death because in other cases, the Product-Limit estimator is not defined beyond the largest observation. Several solutions to this problem are available. First, one can use Efron's tail correction to the Product-Limit estimator (see Practical Note 2 of section 4.2) which changes the largest observed time to a death if it was a censored observation. An estimate of the mean restricted to the interval 0 to $t_{\max}$ is made. A second solution is to estimate the mean restricted to some preassigned interval $[0, \tau]$, where τ is chosen by the investigator to be the longest possible time to which anyone could survive. For either case, the estimated mean restricted to the interval $[0, \tau]$, with τ either the longest observed time or preassigned by the investigator, is given by

$$\hat{\mu}_\tau = \int_0^\tau \hat{S}(t)dt. \tag{4.5.1}$$

The variance of this estimator is

$$\hat{V}[\hat{\mu}_\tau] = \sum_{i=1}^{D}\left[\int_{t_i}^{\tau} \hat{S}(t)dt\right]^2 \frac{d_i}{Y_i(Y_i - d_i)} \tag{4.5.2}$$

A $100(1-\alpha)\%$ confidence interval for the mean is expressed by

$$\hat{\mu}_\tau \pm Z_{1-\alpha/2}\sqrt{\hat{V}[\hat{\mu}_\tau]}. \tag{4.5.3}$$

EXAMPLE 4.1

(continued) Consider estimating the mean survival time for the 6-MP patients based on the Product-Limit estimator presented in Table 4.1. Because the largest observation is censored, an estimate of the mean restricted to 35 weeks will be constructed. The following integrals are needed as intermediate calculations in estimating the variance of our estimate and serve as a convenient bookkeeping method for constructing the estimate of the mean:

$$\int_{23}^{35} \hat{S}(t)dt = 0.448(35 - 23) = 5.376;$$

$$\int_{22}^{35} \hat{S}(t)dt = 5.376 + 0.538(23 - 22) = 5.914;$$

$$\int_{16}^{35} \hat{S}(t)dt = 5.914 + 0.628(22 - 16) = 9.682;$$

$$\int_{13}^{35} \hat{S}(t)dt = 9.682 + 0.690(16 - 13) = 11.752;$$

$$\int_{10}^{35} \hat{S}(t)dt = 11.752 + 0.753(13 - 10) = 14.011;$$

$$\int_{7}^{35} \hat{S}(t)dt = 14.011 + 0.807(10 - 7) = 16.429;$$

$$\int_{6}^{35} \hat{S}(t)dt = 16.429 + 0.857(7 - 6) = 17.286;$$

$$\int_{0}^{35} \hat{S}(t)dt = 17.286 + 1.0(6 - 0) = 23.286.$$

Thus, $\hat{\mu}_{35} = 23.286$ weeks, and

$$\hat{V}[\hat{\mu}_{35}] = \frac{3 \times 17.286^2}{21 \times 18} + \frac{16.429^2}{17 \times 16} + \frac{14.011^2}{15 \times 14} + \frac{11.752^2}{12 \times 11} + \frac{9.682^2}{11 \times 10} + \frac{5.914^2}{7x6} + \frac{5.376^2}{6 \times 5} = 7.993.$$

The standard error of the estimated mean time to relapse is $7.993^{1/2} = 2.827$ weeks.

EXAMPLE 4.2

(continued) Using Efron's tail correction, the estimated mean disease-free survival time for ALL patients is $\hat{\mu}_{2081} = 899.28$ days with a standard error of 150.34 days. A 95% confidence interval for the mean disease-free survival time for ALL patients is $899.28 \pm 1.96(150.34) = (606.61, 1193.95)$ days. Similar calculations for the AML low risk group yields an estimated mean disease-free survival time of $\hat{\mu}_{2569} = 1548.84$ days with a standard error of 150.62 days (95% confidence interval: (1253.62, 1844.07) days.) For the AML high- risk group, $\hat{\mu}_{2640} = 792.31$ days with a standard error of 158.25 days (95% confidence interval: (482.15, 1102.5) days).

Comparison of the duration of the mean disease-free survival time for the three disease categories is complicated by the differences in the largest study times between the groups. To make comparisons which adjust for these differences, the estimated mean, restricted to the interval 0 to 2081 days, is computed for each group. Here, we find the following estimates:

Disease Group	*Mean Restricted to 2081 days*	*Standard Error*	*95% Confidence Interval*
ALL	899.3 days	150.3 days	606.6–1193.9 days
AML low risk	1315.2 days	118.8 days	1082.4–1548.0 days
AML high risk	655.67 days	122.9 days	414.8–896.5 days

Again, these results suggest that AML high risk patients have a lower survival rate than AML low risk patients, whereas ALL patients may be comparable with either of the two AML risk groups.

The Product-Limit estimator can also be used to provide estimates of quantiles of the distribution of the time-to-event distribution. Recall that the pth quantile of a random variable X with survival function $S(x)$, is defined by $x_p = \inf\{t : S(t) \leq 1 - p\}$, that is, x_p is the smallest time at which the survival function is less than or equal to $1 - p$. When $p = 1/2$, x_p is the median time to the event of interest. To estimate x_p, we find the smallest time $\hat{x}_p$ for which the Product-Limit estimator is less than or equal to $1 - p$. That is, $\hat{x}_p = \inf\{t : \hat{S}(t) \leq 1 - p\}$. In practice, the standard error of $\hat{x}_p$ is difficult to compute because it requires an estimate of the density function of X at $\hat{x}_p$ (see Practical Note 3 below). Brookmeyer and Crowley (1982a) have constructed confidence intervals for $\hat{x}_p$, based on a modification of the confidence interval construction for $S(t)$ discussed in section 4.3, which do not require estimating the density function. A $100(1 - \alpha)\%$ confidence interval for x_p, based on the linear confidence interval, is the set of all time points t which satisfy the following condition:

$$-Z_{1-\alpha/2} \leq \frac{\hat{S}(t) - (1-p)}{\hat{V}^{1/2}[\hat{S}(t)]} \leq Z_{1-\alpha/2}. \tag{4.5.4}$$

The $100(1-\alpha)\%$ confidence interval for x_p based on the log-transformed interval is the set of all points t which satisfy the condition:

$$-Z_{1-\alpha/2} \leq \frac{[\ln\{-\ln[\hat{S}(t)]\} - \ln\{-\ln[1-p]\}][\hat{S}(t)\ln[\hat{S}(t)]}{\hat{V}^{1/2}[\hat{S}(t)]} \leq Z_{1-\alpha/2}. \tag{4.5.5}$$

The $100(1-\alpha)\%$ confidence interval for x_p based on the arcsine-square root transformation is given by

$$-Z_{1-\alpha/2} \leq \frac{2\{\text{arcsine}[\sqrt{\hat{S}(t)}] - \text{arcsine}[\sqrt{(1-p)}]\}[\hat{S}(t)(1-\hat{S}(t))]^{1/2}}{\hat{V}^{1/2}[\hat{S}(t)]} \leq Z_{1-\alpha/2}. \tag{4.5.6}$$

EXAMPLE 4.2 *(continued)* We shall estimate the median disease-free survival time for the ALL group. From Table 4.3 we see that $\hat{S}(383) = 0.5217 > 0.5$

TABLE 4.7
Construction of a 95% Confidence Interval for the Median

t_i	$\hat{S}(t_i)$	$\sqrt{\hat{V}[\hat{S}(t_i)]}$	*Linear* (4.5.4)	*Log* (4.5.5)	*Arcsine* (4.5.6)
1	0.9737	0.0260	18.242	3.258	7.674
55	0.9474	0.0362	12.350	3.607	6.829
74	0.9211	0.0437	9.625	3.691	6.172
86	0.8947	0.0498	7.929	3.657	5.609
104	0.8684	0.0548	6.719	3.557	5.107
107	0.8421	0.0592	5.783	3.412	4.645
109	0.8158	0.0629	5.022	3.236	4.214
110	0.7895	0.0661	4.377	3.036	3.806
122	0.7368	0.0714	3.316	2.582	3.042
129	0.7105	0.0736	2.862	2.334	2.679
172	0.6842	0.0754	2.443	2.074	2.326
192	0.6579	0.0770	2.052	1.804	1.981
194	0.6316	0.0783	1.681	1.524	1.642
230	0.6041	0.0795	1.309	1.220	1.290
276	0.5767	0.0805	0.952	0.909	0.945
332	0.5492	0.0812	0.606	0.590	0.604
383	0.5217	0.0817	0.266	0.263	0.266
418	0.4943	0.0819	−0.070	−0.070	−0.070
468	0.4668	0.0818	−0.406	−0.411	−0.405
487	0.4394	0.0815	−0.744	−0.759	−0.741
526	0.4119	0.0809	−1.090	−1.114	−1.078
609	0.3825	0.0803	−1.464	−1.497	−1.437
662	0.3531	0.0793	−1.853	−1.886	−1.798
2081	0.3531	0.0793	−1.853	−1.886	−1.798

and $\hat{S}(418) = 0.4943 \leq 0.5$, so $\hat{x}_{0.5} = 418$ days. To construct 95% confidence intervals for the median, we complete Table 4.7. To illustrate the calculations which enter into construction of this Table, consider the first row. Here the entry in the fourth column is the middle term in (4.5.4), namely, $(0.9737 - 0.5)/0.0260 = 18.242$. The entry in the fifth column is the middle term in (4.5.5), namely,

$$([\ln(-\ln(0.9737)) - \ln(-\ln(0.5))]\{0.9737\ln[0.9737]\}/0.0260) = 3.258,$$

and the entry in the last column is the middle term in (4.5.6), namely, $2[\text{arcsine}(\sqrt{0.9737}) - \text{arcsine}(\sqrt{0.5})][0.9737(1 - 0.9737)]^{1/2}/0.0260 = 7.674$. To find the linear 95% confidence interval, we find all those values of t_i which have a value, in column four between -1.96 and 1.96. Thus the 95% linear confidence interval for $x_{0.5}$ is $x_{0.05} > 194$ days. The upper limit of this interval is undetermined because (4.5.4)

never drops below -1.96 due to the heavy censoring. Based on the log transformation, a 95% confidence interval for x_p is $x_{0.05} > 192$ days. The interval based on the arcsine-transformed interval is $x_{0.05} > 194$ days.

Similar calculations for the two groups of AML patients show that the median disease-free survival time, for the low risk group, is 2204 days and, for the high risk group, is 183 days. For the low risk group, the lower end points of the 95% confidence intervals for the median disease-free survival time are 704 days, based on the linear approximation and 641 days based on either the log or arcsine transformation. For the high risk group, the 95% confidence intervals for the median are (115, 363) days for the linear and arcsine-square root transformed intervals and (113, 363), based on the log-transformed interval.

Practical Notes

1. If there is no censoring, then, the estimator of the mean time to death reduces to the sample mean. In addition, if there are no ties, then the estimated variance of the mean estimate reduces to the sample variance divided by n.
2. Alternate estimators of the mean survival time can be found by finding the area under one of the tail-corrected Product-Limit estimators discussed in Practical Note 2 of section 4.2.
3. An estimator of the large sample variance of the estimator of the pth percentile is given by $\hat{V}[\hat{x}_p] = \frac{\hat{V}[\hat{S}(x_p)]}{\hat{f}(x_p)^2}$, where $\hat{f}(x_p)$ is an estimate of the density function at the pth percentile. A crude estimate of $\hat{f}(t)$ is $\frac{\hat{S}(t-b)-\hat{S}(t+b)}{2b}$ based on a uniform kernel density estimate. Here, b is some small number.
4. Most major statistical packages provide an estimate of the mean lifetime. When the largest observation is censored, one must carefully examine the range over which the mean is computed.

Theoretical Notes

1. The asymptotic properties of the estimators of the mean and pth quantile follow directly from the weak convergence of the Product-Limit estimator. Details can be found in Andersen et al. (1993).
2. Details of constructing the confidence interval for median survival are found in Brookmeyer and Crowley (1982a) who also present a Monte Carlo study of the performance of the linear interval.

4.6 Estimators of the Survival Function for Left-Truncated and Right-Censored Data

The estimators and confidence intervals presented in sections 4.2–4.5 were based on right-censored samples. In this section, we shall show how these statistics can be modified to handle left-truncated and right-censored data. Here, we have associated, with the jth individual, a random age L_j at which he/she enters the study and a time T_j at which he/she either dies or is censored. As in the case of right-censored data, define $t_1 < t_2 < \cdots < t_D$ as the distinct death times and let d_i be the number of individuals who experience the event of interest at time t_i. The remaining quantity needed to compute the statistics in the previous sections is the number of individuals who are at risk of experiencing the event of interest at time t_i, namely Y_i. For right-censored data, this quantity was the number of individuals on study at time 0 with a study time of at least t_i. For left-truncated data, we redefine Y_i as the number of individuals who entered the study prior to time t_i and who have a study time of at least t_i, that is, Y_i is the number of individuals with $L_j < t_i \leq T_j$.

Using Y_i as redefined for left-truncated data, all of the estimation procedures defined in sections 4.2–4.4 are now applicable. However, one must take care in interpreting these statistics. For example, the Product-Limit estimator of the survival function at a time t is now an estimator of the probability of survival beyond t, conditional on survival to the smallest of the entry times $L, Pr[X > t \mid X \geq L] = S(t)/S(L)$. Similarly the Nelson–Aalen statistic estimates the integral of the hazard rate over the interval L to t. Note that the slope of the Nelson–Aalen estimator still provides an estimator of the unconditional hazard rate.

Some care in directly applying these estimators is needed. For left-truncated data, it is possible for the number at risk to be quite small for small values of t_i. If, for some t_i, Y_i and d_i are equal, then, the Product-Limit estimator will be zero for all t beyond this point, even though we are observing survivors and deaths beyond this point. In such cases, it is common to estimate the survival function conditional on survival to a time where this will not happen by considering only those death times beyond this point. This is illustrated in the following example.

EXAMPLE 4.3

To illustrate how the statistics developed in the previous sections can be applied to left-truncated data, consider the Channing House data described in section 1.16. The data is found in Table D.5 of Appendix D. Here the truncation times are the ages, in months, at which individuals

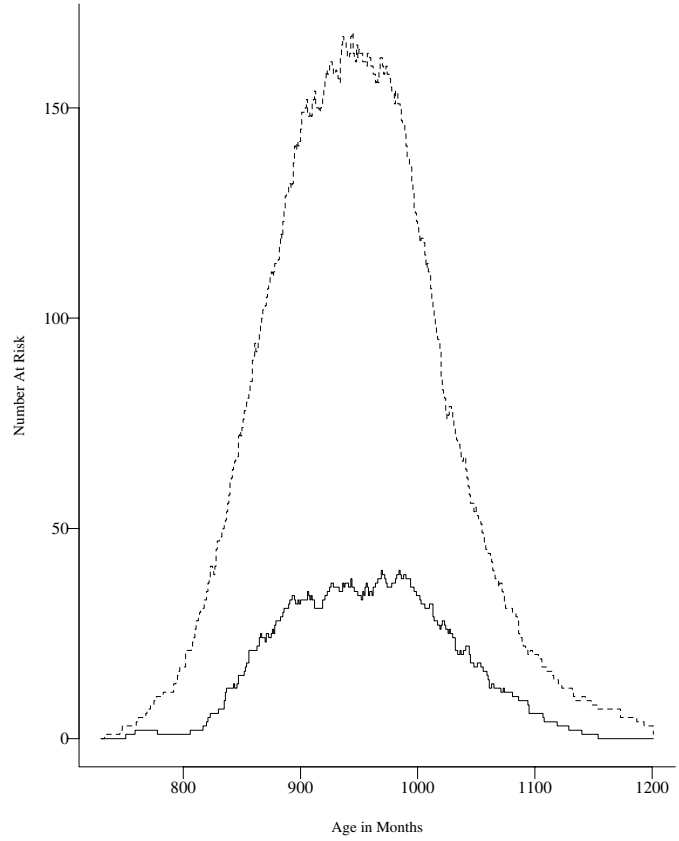

Figure 4.10 *Number at risk as a function of age for the 97 males (———) and the 365 females (-----) in the Channing house data set*

entered the community. We shall focus on estimating the conditional survival function.

Figure 4.10 shows the number of individuals at risk as a function of the age at which individuals die for both males and females. Note that the number at risk initially increases as more individuals enter into the study cohort and that this number decreases for later ages as individuals die or are censored.

Consider the data on males. Here the risk set is empty until 751 months when one individual enters the risk set. At 759 months, a second individual enters the risk set. These two individuals die at 777 and 781 months. A third individual enters the risk set at 782 months. Computing the Product-Limit estimator of $S(t)$ directly by (4.2.1) based on this data would yield an estimate of $\hat{S}(t) = 1$ for $t < 777$, $\hat{S}(t) = 1/2$ for $777 \leq t < 781$, and $\hat{S}(t) = 0$ for $t \geq 781$. This estimate has little

meaning since the majority of the males in the study clearly survive beyond 781 months.

Rather than estimating the unconditional survival function, we estimate the conditional probability of surviving beyond age t, given survival to age a. We estimate $S_a(t) = Pr[X > t \mid X \geq a]$ by considering only those deaths that occur after age a, that is,

$$\hat{S}_a(t) = \prod_{a \leq t_i \leq t} \left[1 - \frac{d_i}{Y_i}\right], t \geq a. \tag{4.6.1}$$

Similarly for Greenwood's formula (4.2.2) or for the Nelson–Aalen estimator (4.2.3), only deaths beyond a are considered.

Figure 4.11 shows the estimated probability of surviving beyond age t, given survival to 68 or 80 years for both males and females.

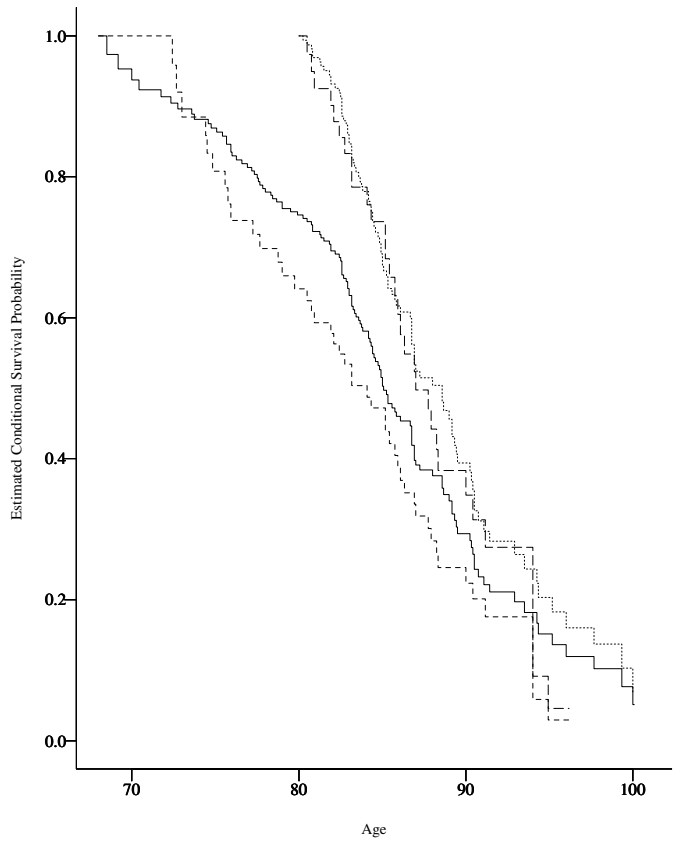

Figure 4.11 *Estimated conditional survival functions for Channing house residents. 68 year old females (———); 80 year old females (------); 68 year old males (— — —); 80 year old males (—— ——).*

As in the unconditional Product-Limit estimator, the estimates are computed only over the range where $Y_i > 0$. These estimates could be extended beyond this time by the methods discussed in Practical Note 3 of section 4.2.

Practical Notes

1. A key assumption needed for making an inference with left-truncated data is the notion of 'independent truncation', that is, the Product-Limit estimator is a maximum likelihood estimator of the survival function if

$$\frac{Pr[X = x \mid L = l, X > L]}{Pr[X \geq x \mid L = l, X > L]} = \frac{Pr[X = x]}{Pr[X \geq x]} = h(x),$$

   the hazard rate of X. Tsai (1990) provides a formal test of this hypothesis which is valid under independent censoring. See Keiding (1992) for further discussion of this point and further examples.
2. When the data is left truncated and right censored, the Product-Limit estimator may have a relatively large variance for small t, where the risk sets are small, and also for large t. This early instability in the point estimator of the survival function may propagate throughout the entire curve. Lai and Ying (1991) suggest a solution to this problem by a slight modification of the Product-Limit estimators where deaths are ignored when the risk set is small. Their estimator is given by

$$\tilde{S}(t) = \prod_{t_i \leq t} \left\{ 1 - \frac{d_i}{Y_i} I[Y_i \geq cn^{\alpha}] \right\},$$

   where I is the indicator of the set A, n is the sample size, and $c > 0$, $0 < \alpha < 1$ are constants. This estimator is asymptotically equivalent to the usual product limit estimator.

Theoretical Note

1. The derivation of the Product-Limit estimator and the Nelson–Aalen estimator follows directly from the theory of counting processes as presented in section 3.6 with the modified definition of $Y(t)$ as discussed in Practical Note 2 of that section.

4.7 Summary Curves for Competing Risks

The summary survival curves presented in sections 4.2–4.6 are based on the assumption that the event and censoring times are independent. In the case of competing risks data, as discussed in section 2.7, this untestable assumption may be suspect. In this section we present three techniques for summarizing competing risks data.

To help in understanding the difference between the three estimators and their interpretation, consider the bone marrow transplant study discussed in section 1.3. In earlier sections of this chapter we considered estimation of the survival function for the time to treatment failure. Recall that treatment failure is defined as death in remission or relapse, whichever comes first. Here death in remission and relapse are competing risks and we are interested in summary curves that tell us how the likelihood of these events develops over time. Occurrence of one of the events precludes occurrence of the other event.

The first estimator which is commonly used is the complement of the Kaplan-Meier estimator. Here occurrences of the other event are treated as censored observations. For example, the estimated probability of relapsing before time t is one minus the Kaplan-Meier estimator of relapse obtained by treating occurrences of relapse as events and occurrences of death before relapse as censored observations. This estimator is an attempt to estimate the probability of relapsing before time t. It can be interpreted as the probability of relapse by time t if the risk of non-relapse death was removed. It is the probability of relapse in a hypothetical world where it is impossible for patients to die in remission. Reversing the roles of death in remission and relapse yields the treatment-related mortality or death in remission probability. Here this is an estimate of death in the world where relapse is not possible. These are rarely the probabilities of clinical interest and we cannot recommend the use of this estimator.

The second estimator is the cumulative incidence function. This estimator is constructed as follows. Let $t_1 < t_2 < \cdots < t_K$ be the distinct times where one of the competing risks occurs. At time t_i let Y_i be the number of subjects at risk, r_i be the number of subjects with an occurrence of the event of interest at this time, and d_i be the number of subjects with an occurrence of any of the other events of interest at this time. Note that $d_i + r_i$ is the number of subjects with an occurrence of any one of the competing risks at this time. Independent random censoring due to a patient being lost to follow-up is not counted here as one of the competing risks and affects only the value of Y_i. The cumulative incidence function is defined by

$$\mathrm{CI}(t) = \begin{cases} 0 & \text{if } t \le t_1 \\ \displaystyle\sum_{t_i \le t} \left\{ \prod_{j=1}^{i-1} \frac{1 - [d_j + r_j]}{Y_j} \right\} \frac{r_i}{Y_i} & \text{if } t_1 \le t \end{cases}. \tag{4.7.1}$$

Note that for $t \ge t_1$ the cumulative incidence function is

$$\mathrm{CI}(t) = \sum_{t_i \le t} \hat{S}(t_{i^-}) \frac{r_i}{Y_i}$$

where $\hat{S}(t_{i^-})$ is the Kaplan-Meier estimator, evaluated at just before t_i, obtained by treating any one of the competing risks as an event. The cumulative incidence function estimates the probability that the event of interest occurs before time t and that it occurs before any of the competing causes of failure. It is the estimate of the probability of the event of interest in the real world where a subject may fail from any of the competing causes of failure. For example, the relapse cumulative incidence is the chance a patient will have relapsed in the interval 0 to t in the real world where they may die in remission. The treatment related mortality cumulative incidence is the chance of death before relapse in the real world. Note that the sum of the cumulative incidences for all the competing risks is $1 - \hat{S}(t)$, which in the bone marrow transplant example is the complement of the treatment failure Kaplan-Meier estimate found in section 4.2.

The variance of the cumulative incidence is estimated by

$$V[\mathrm{CI}(t)] = \sum_{t_i \le t} \hat{S}(t_i)^2 \left\{ [\mathrm{CI}(t) - \mathrm{CI}(t_i)]^2 \frac{r_i + d_i}{Y_i^2} + [1 - 2(\mathrm{CI}(t) - \mathrm{CI}(t_i))] \frac{r_i}{Y_i^2} \right\}. \tag{4.7.2}$$

Confidence pointwise $(1 - \alpha)$ 100% confidence intervals for the cumulative incidence are given by $\mathrm{CI}(t) \pm Z_{1-\alpha/2} V[\mathrm{CI}(t)]^{1/2}$.

The third probability used to summarize competing risks data is the conditional probability function for the competing risk. For a particular risk, K, let $\mathrm{CI}_K(t)$ and $\mathrm{CI}_{K^c}(t)$ be the cumulative incidence functions for risk K and for all other risks lumped together, respectively. Then the conditional probability function is defined by

$$\mathrm{CP}_K(t) = \frac{\mathrm{CI}_K(t)}{1 - \mathrm{CI}_{K^c}(t)}. \tag{4.7.3}$$

The variance of this statistic is estimated by

$$V[\mathrm{CP}_K(t)] = \frac{\hat{S}(t-)^2}{\{1 - \mathrm{CI}_{K^c}(t)\}^4} \sum_{t_i \le t} \frac{[1 - \mathrm{CI}_{K^c}(t_i)]^2 r_i + \mathrm{CI}_K(t_i)^2 d_i}{Y_i^2}. \tag{4.7.4}$$

The conditional probability is an estimate of the conditional probability of event K's occurring by t given that none of the other causes have occurred by t. In the bone marrow transplantation example the conditional probability of relapse is the probability of relapsing before time t given the patient is not dead from other causes prior to t. It is the probability of relapsing among survivors who have not died from non-relapse-related toxicities.

To understand these probabilities better, consider a hypothetical bone marrow transplant experiment involving 100 patients. Suppose that there is no independent censoring and at one year after transplant 10 patients have relapsed and 30 patients have died in remission. When there is no censoring the cumulative incidence reduces to the cumulative number of events of the given type divided by n so the relapse cumulative incidence is 10/100 and the death in remission cumulative incidence is 30/100. The death in remission incidence is clearly interpreted as the proportion of patients who died in complete remission

TABLE 4.8
Estimates of Relapse and Death in Remission (TRM) for ALL Patients

t_i	d_i	r_i	Y_i	*TRM 1-KME*	*Relapse 1-KME*	*TRM CI*	*Relapse CI*	*TRM CP*	*Relapse CP*
1	1	0	38	0.0263	0.0000	0.0263	0.0000	0.0263	0.0000
55	0	1	37	0.0263	0.0270	0.0263	0.0263	0.0270	0.0270
74	0	1	36	0.0263	0.0541	0.0263	0.0526	0.0278	0.0541
86	1	0	35	0.0541	0.0541	0.0526	0.0526	0.0556	0.0556
104	0	1	34	0.0541	0.0819	0.0526	0.0789	0.0571	0.0833
107	1	0	33	0.0828	0.0819	0.0789	0.0789	0.0857	0.0857
109	0	1	32	0.0828	0.1106	0.0789	0.1053	0.0882	0.1143
110	0	1	31	0.0828	0.1393	0.0789	0.1316	0.0909	0.1429
122	1	1	30	0.1134	0.1680	0.1053	0.1579	0.1250	0.1765
129	0	1	28	0.1134	0.1977	0.1053	0.1842	0.1290	0.2059
172	1	0	27	0.1462	0.1977	0.1316	0.1842	0.1613	0.2121
192	0	1	26	0.1462	0.2285	0.1316	0.2105	0.1667	0.2424
194	1	0	25	0.1804	0.2285	0.1579	0.2105	0.2000	0.2500
230	0	1	23	0.1804	0.2621	0.1579	0.2380	0.2072	0.2826
276	1	0	22	0.2176	0.2621	0.1854	0.2380	0.2432	0.2921
332	0	1	21	0.2549	0.2621	0.2128	0.2380	0.2793	0.3023
383	0	1	20	0.2549	0.2990	0.2128	0.2654	0.2897	0.3372
418	1	0	19	0.2941	0.2990	0.2403	0.2654	0.3271	0.3494
466	1	0	18	0.3333	0.2990	0.2677	0.2654	0.3645	0.3625
487	1	0	17	0.3725	0.2990	0.2952	0.2654	0.4019	0.3766
526	1	0	16	0.4117	0.2990	0.3227	0.2654	0.4393	0.3919
609	0	1	14	0.4117	0.3490	0.3227	0.2949	0.4576	0.4353
662	0	1	13	0.4117	0.3991	0.3227	0.3243	0.4775	0.4788

prior to one year. The conditional probabilities estimates are 10/70 for relapse and 30/90 for death in remission. Here the death in remission probability is estimated by the number who die in remission divided by the number who could have died in remission which is the number at risk at one year who have yet to relapse. The complement of the Kaplan-Meier estimate depends on the pattern of occurrences of deaths and relapses. If all deaths occur before the first relapse then the relapse probability is 10/70 while if all the relapses occurred before the first death we get an estimate of 10/100. Any value between these two extremes is possible. Clearly this estimate has no meaningful interpretation.

EXAMPLE 4.2

(continued) We consider the data on the 38 patients with ALL given a transplant and examine the three probabilities for relapse and for death in remission (TRM). Table 4.8 provides the estimates for the three probabilities. The estimated standard error for the relapse cumulative incidence at 1 year is 0.069 so an approximate 95% confidence interval for the probability of relapsing before death is $0.238 \pm 1.96 \times 0.069 =$

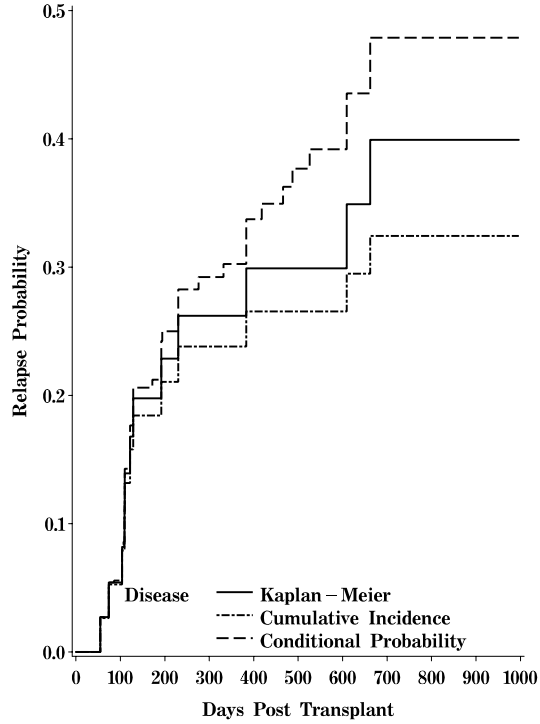

Figure 4.12 *Comparison estimated probability of relapse for ALL patients. Complement of Kaplan-Meier* (——), *cumulative incidence* (−·−·), *conditional probability* (- - -)

(0.103, 0.373). The estimated conditional probability of relapse at 1 year was 0.302 with a standard error of 0.087. A 95% confidence interval for the conditional probability of relapse is $0.302 \pm 1.96 \times 0.087 = (0.131, 0.473)$.

Figures 4.12 and 4.13 show the estimated probabilities for relapse and death in remission, respectively. Note that the conditional probability curve changes value at the occurrence of either of the two competing risks. The probabilities have the characteristic property of the conditional probability estimate being the largest and the cumulative incidence estimate the smallest.

It is important that summary curves for all the competing risks be presented since changes in the likelihood of one event cause changes in the probabilities for the other events. A nice summary curve is shown in Figure 4.14. Here we plot the relapse cumulative incidence and the sum of the relapse and death in remission cumulative incidences. The complement of the sum of the two cumulative incidences is the disease free survival probability found in section 4.2. At a given time the height of the first curve is the probability of relapsing, the distance between

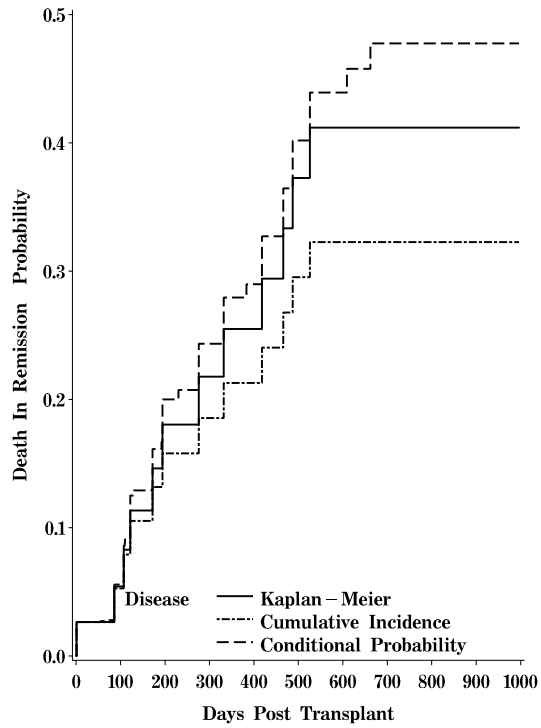

Figure 4.13 *Comparison estimated probability of death in remission for ALL patients. Complement of Kaplan-Meier* (——), *cumulative incidence* (−·−·), *conditional probability* (- - -)

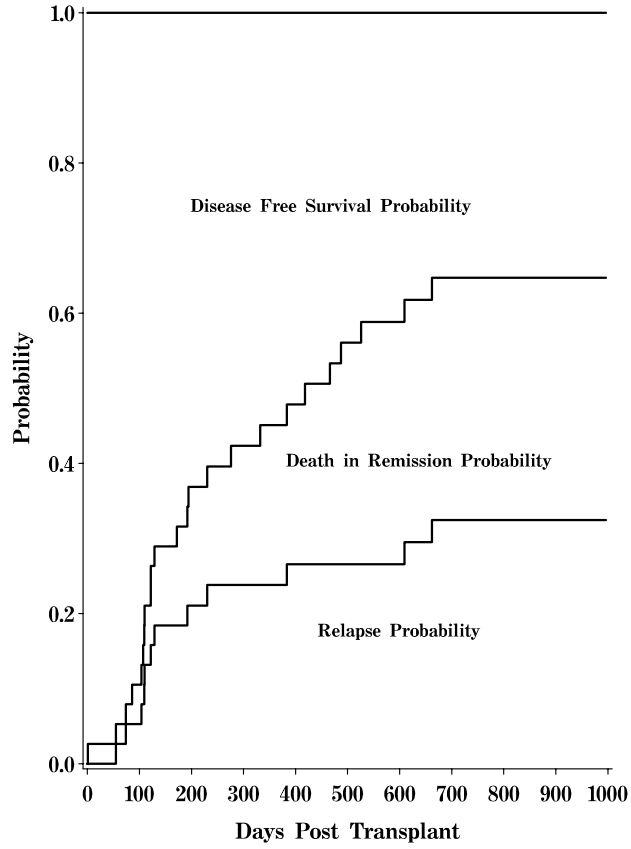

Figure 4.14 *Interaction between the relapse and death in remission*

the first and second curves the probability of death in remission, and the distance of the second curve from 1.0 the disease free survival function. For example, at 400 days the relapse probability is 0.2654, the death in remission probability is $0.4982 - 0.2654 = 0.2328$, and the disease-free survival function is $1 - 0.4982 = 0.5018$. This graph allows us dynamically to access the relationship between the competing risks.

Theoretical Notes

1. Suppose we have two competing risks X and Y and let $T = \min(X, Y)$ and $I = 1$ if $X < Y$, 0 if $X > Y$. The cause-specific hazard rate for X is

$$\lambda_X(t) = P[t \leq X < t + \Delta t \mid \min(X, Y) > t]\,\Delta t.$$

   The Kaplan-Meier estimator obtained by treating times with $I = 0$ as censored observations provides a consistent estimator of

$\exp\{-\int_0^t \lambda_X(u)\,du\}$. This quantity has no interpretation as a probability.

2. The cumulative incidence estimator was first proposed by Kalbfleisch and Prentice (1980). The estimator can be derived using techniques described in Andersen et al. (1993) as a special case of a more general theory for product-limit estimators for the transitions of a non-homogeneous Markov process.
3. Pepe and Mori (1993), Pepe et al. (1993), and Gooley et al. (1999) provide a nice discussion of these three estimates and present alternative derivations of the variance estimates.

Practical Note

1. A SAS macro to compute the cumulative incidence curves can be found on our web site.

4.8 Exercises

4.1 In section 1.11 we discussed a study of the effect of ploidy on the survival of patients with cancer of the tongue. Using the data on aneuploid tumors found in Table 1.6.

(a) Estimate the survival function at one (12 months) and five years (60 months) after transplant. Find the standard errors for your estimates.

(b) Estimate the cumulative hazard rate, $H(t)$, at 60 months. Find the standard error of $\hat{H}(t)$. Estimate $S(60)$ by $\exp\{-\hat{H}(t)\}$ and compare to your estimate in part a.

(c) Find a 95% linear confidence interval for $S(60)$.

(d) Find a 95% log-transformed confidence interval for $S(60)$.

(e) Find a 95% arcsine-square root confidence interval for $S(60)$.

(f) Using the log transformation find a 95% EP confidence band for the survival function over the range three years to six years (i.e., 36–72 months).

(g) Using the log transformation find a 95% Hall-Wellner confidence band for the survival function over the range three years to six years (i.e., 36–72 months).

(h) Estimate the mean survival time restricted to 400 months. Also provide a 95% confidence interval for the restricted mean survival time.

(i) Estimate the median time to death and find a 95% confidence interval for the median survival time based on a linear confidence interval.

4.2 Using the data reported in section 1.3, find the quantities specified below for the AML low risk and AML high risk groups. Note that most of these quantities are worked out in detail in Example 4.2 and its continuations for the ALL group.

(a) Estimate the survival functions and their standard errors for the AML low risk and AML high risk groups.

(b) Estimate the cumulative hazard rates and their standard errors for the AML low risk and AML high risk groups.

(c) Provide a crude estimate of the hazard rates for each group based on the estimates obtained in (b).

(d) Estimate the mean time to death and find 95% confidence intervals for the mean survival time for both the AML low risk and AML high risk groups. (Answers are given in section 4.5.)

(e) Work out estimates of the median time to death and find 95% confidence intervals for the median survival time for both the AML low risk and AML high risk groups using the linear, log-transformed, and arcsine formulas. (Answers are given in section 4.5.)

(f) Find 95% confidence intervals for the survival functions at 300 days post-transplant for both the AML low risk and AML high risk groups using the log- and arcsine-transformed formulas.

(g) Find 95% EP confidence bands for the survival functions over the range 100–400 days post-transplant for both the AML low risk and AML high risk groups using the linear, log-transformed, and arcsine-transformed formulas.

(h) Find 95% HW confidence bands for the survival functions over the range 100–400 days post-transplant for both the AML low risk and AML high risk groups using the linear, log-transformed, and arcsine-transformed formulas.

(i) Based on the results above and those discussed in Example 4.2 and its continuations, how do the survival experiences of the ALL, AML low risk, and AML high risk groups compare?

4.3 The following table contains data on the survival times of 25 patients with inoperative lung cancer entered on a study between November 1, 1979, and December 23, 1979. Complete follow-up was obtained on all patients so that the exact date of death was known. The study had one interim analysis conducted on March 31, 1980, by which time only 13 patients had died.

(a) Estimate the survival function based on the available sample information at the time of the interim analysis on 3/31/80. Provide the standard error of your estimate.

(b) Use the Brown, Hollandar, and Kowar technique (Practical Note 2 of section 4.1) to complete the right-hand tail of the product-limit estimate found in part a.

Patient	*Date of Diagnosis*	*Date of Death*	*Days to death*	*Days to 3/31/80(Status)*
1	1/11/79	5/30/79	139	139(Dead)
2	1/23/79	1/21/80	363	363(Dead)
3	2/15/79	8/27/79	193	193(Dead)
4	3/7/79	11/10/79	248	248(Dead)
5	3/12/79	4/8/79	27	27(Dead)
6	3/25/79	10/21/79	210	210(Dead)
7	4/4/79	8/16/79	134	134(Dead)
8	4/30/79	11/19/79	203	203(Dead)
9	5/16/79	5/9/81	724	320 (Alive)
10	5/26/79	7/15/79	50	50(Dead)
11	5/30/79	10/22/80	511	306(Alive)
12	6/3/79	6/25/79	22	22(Dead)
13	6/15/79	12/27/80	561	290(Alive)
14	6/29/79	1/29/81	580	276(Alive)
15	7/1/79	11/14/79	136	136(Dead)
16	8/13/79	6/16/80	308	231(Alive)
17	8/27/79	4/7/80	224	217(Alive)
18	9/15/79	1/9/81	482	198(Alive)
19	9/27/79	4/5/80	191	186(Alive)
20	10/11/79	3/3/80	144	144(Dead)
21	11/17/79	1/24/80	68	68(Dead)
22	11/21/79	10/4/81	683	131(Alive)
23	12/1/79	8/13/80	256	121(Alive)
24	12/14/79	2/27/81	441	108(Alive)
25	12/23/79	4/2/80	101	99(Alive)

(c) Compute the estimate of the survival function and an estimate of its standard error using the complete follow-up on each patient. Compare this estimate to that found in part a.

(d) Estimate the mean time to death restricted to 683 days based on the product-limit estimator found in part c.

(e) Estimate the mean time to death by finding the area under the survival curve found in part c. Find the standard error of your estimate.

(f) Compute the usual estimate of the time to death based on complete follow-up data by finding the arithmetic mean of the complete follow-up data. Find the standard error of this estimate in the usual way as the sample standard deviation divided by the square root of the sample size. Compare your answers to those obtained in part e.

4.4 In section 1.4 the times to first exit site infection (in months) of patients with renal insufficiency was reported. In the study 43 patients had a surgically placed catheter (Group 1) and 76 patients had a percutaneous placement of their catheter (Group 0).

(a) For each group plot the estimated survival function. Which technique seems better in delaying the time to infection?

(b) Estimate the cumulative hazard rate for each group of patients. Provide a crude estimate of the hazard rate at 5 months after placement of the catheter in each group.

(c) Find a 95% confidence interval for the mean time to first exit site infection restricted to 36 months for both groups.

4.5 Using the survival times of 59 black females given a kidney transplant at the OSU transplant center discussed in section 1.7—

(a) Estimate the distribution of the time to death, measured from transplant, for black female kidney transplant patients. Provide the standard error of the estimated survival function.

(b) Find a 95% confidence interval, based on the linear transformation, for the probability a black female will survive at least 12 months (365 days) after transplantation.

(c) Repeat b using the log-transformed confidence interval.

(d) Repeat c using the arcsine-transformed confidence interval. Compare the intervals found in parts c–e.

4.6 In section 1.6 a study is described to evaluate a protocol change in disinfectant practice in a large midwestern university medical center. Control of infection is the primary concern for the 155 patients entered into the burn unit with varying degrees of burns. The outcome variable is the time until infection from admission to the unit. Censoring variables are discharge from the hospital without an infection or death without an infection. Eighty-four patients were in the group which had chlorhexidine as the disinfectant and 72 patients received the routine disinfectant povidone-iodine.

(a) Estimate the survival (infection-free) functions and their standard errors for the chlorhexidine and povidone-iodine groups.

(b) Estimate the cumulative hazard rates and their standard errors for the chlorhexidine and povidone-iodine groups. Plot these estimates. Does it appear that the two cumulative hazard rates are proportional to each other?

(c) Provide estimates of the median time to infection and find 95% confidence intervals for the median time to infection for both the chlorhexidine and povidone-iodine groups using the linear, log-transformed, and arcsine formulas.

(d) Find 95% confidence intervals for the survival (infection-free) functions at 10 days postadmission for both the chlorhexidine and povidone-iodine groups using the log transformed and arcsine transformed formulas.

(e) Find 95% confidence bands for the infection-free functions over the range 8–20 days postinfection for both the chlorhexidine and povidone-

iodine groups using the linear, log transformed, and arcsine transformed formulas.

(f) Find 95% HW confidence bands for the infection-free functions over the range 8–20 days postinfection for both the chlorhexidine and povidone-iodine.

(g) Based on the results above, how does the infection experience of the chlorhexidine and povidone-iodine groups compare?

4.7 Consider a hypothetical study of the mortality experience of diabetics. Thirty diabetic subjects are recruited at a clinic and followed until death or the end of the study. The subject's age at entry into the study and their age at the end of study or death are given in the table below. Of interest is estimating the survival curve for a 60- or for a 70-year-old diabetic.

(a) Since the diabetics needed to survive long enough from birth until the study began, the data is left truncated. Construct a table showing the number of subjects at risk, Y, as a function of age.

(b) Estimate the conditional survival function for the age of death of a diabetic patient who has survived to age 60.

(c) Estimate the conditional survival function for the age of death of a diabetic patient who has survived to age 70.

(d) Suppose an investigator incorrectly ignored the left truncation and simply treated the data as right censored. Repeat parts a–c.

Entry Age	*Exit Age*	*Death Indicator*	*Entry Age*	*Exit Age*	*Death Indicator*
58	60	1	67	70	1
58	63	1	67	77	1
59	69	0	67	69	1
60	62	1	68	72	1
60	65	1	69	79	0
61	72	0	69	72	1
61	69	0	69	70	1
62	73	0	70	76	0
62	66	1	70	71	1
62	65	1	70	78	0
63	68	1	71	79	0
63	74	0	72	76	1
64	71	1	72	73	1
66	68	1	73	80	0
66	69	1	73	74	1

4.8 Table 1.7 reports the results of a study on the survival times of patients admitted to a psychiatric hospital. In this data set patients were admitted to the hospital at a random age and followed until death or the end of the study. Let X be the patient's age at death. Note that the data we

have on X is left truncated by the patient's age at entry into the hospital and right censored by the end of the study.

(a) Plot the number at risk, Y_i, as a function of age.

(b) Estimate the conditional survival function for a psychiatric patient who has survived to age 30 without entering a psychiatric hospital.

4.9 Hoel and Walburg (1972) report results of an experiment to study the effects of radiation on life lengths of mice. Mice were given a dose of 300 rads of radiation at 5–6 weeks of age and followed to death. At death each mouse was necropsied to determine if the cause of death was thymic lymphoma, reticulum cell sarcoma, or another cause. The ages of the mice at death are shown below:

Cause of Death	*Age at Death (Days)*
Thymic lymphoma	158, 192, 193, 194, 195, 202, 212, 215, 229, 230, 237, 240, 244, 247, 259, 300, 301, 337, 415, 444, 485, 496, 529, 537, 624, 707, 800
Reticulum cell sarcoma	430, 590, 606, 638, 655, 679, 691, 693, 696, 747, 752, 760, 778, 821, 986
Other causes	136, 246, 255, 376, 421, 565, 616, 617, 652, 655, 658, 660, 662, 675, 681, 734, 736, 737, 757, 769, 777, 801, 807, 825, 855, 857, 864, 868, 870, 873, 882, 895, 910, 934, 942, 1,015, 1,019

(a) For each of the three competing risks estimate the cumulative incidence function at 200, 300, . . . , 1,000 days by considering the two other risks as a single competing risk.

(b) Show that the sum of the three cumulative incidence functions found in part a is equal to the Kaplan-Meier estimate of the overall survival function for this set of data.

(c) Repeat part a using the complement of the marginal Kaplan-Meier estimates. What are the quantities estimating and how different from the results found in part a are these estimates?

(d) Compute the conditional probability function for thymic lymphoma at 500 and 800 days. What are the quantities estimating?

4.10 Using the data reported in section 1.3 for the AML low risk and AML high risk groups, find the following quantities for the two competing risks of relapse and death:

(a) The estimated cumulative incidence at one year.

(b) The standard errors of the two estimates in part a.

(c) The estimated conditional probabilities of relapse and of death in remission.

(d) The standard errors of the probabilities found in part c.

(e) Graphically express the development of relapse and death in remission for these two disease groups.