8 Recurrent Event Survival Analysis

D.G. Kleinbaum and M. Klein, *Survival Analysis: A Self-Learning Text, Third Edition*,
Statistics for Biology and Health, DOI 10.1007/978-1-4419-6646-9_8,

Introduction

This chapter considers outcome events that may occur more than once over the follow-up time for a given subject. Such events are called "recurrent events." Modeling this type of data can be carried out using a Cox PH model with the data layout constructed so that each subject has a line of data corresponding to each recurrent event. A variation of this approach uses a stratified Cox PH model, which stratifies on the order in which recurrent events occur. Regardless of which approach is used, the researcher should consider adjusting the variances of estimated model coefficients for the likely correlation among recurrent events on the same subject. Such adjusted variance estimates are called "robust variance estimates." A parametric approach for analyzing recurrent event data that includes a frailty component (introduced in Chapter 7) is also described and illustrated.

Abbreviated Outline

The outline below gives the user a preview of the material to be covered by the presentation. A detailed outline for review purposes follows the presentation.

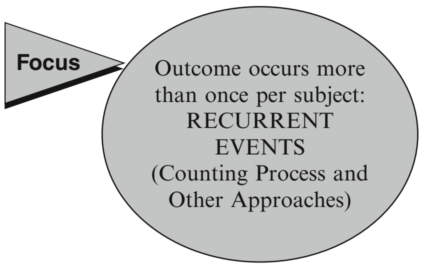
Objectives

Upon completing this chapter, the learner should be able to:

1. State or recognize examples of recurrent event data.
2. State or recognize the form of the data layout used for the counting process approach for analyzing correlated data.
3. Given recurrent event data, outline the steps needed to analyze such data using the counting process approach.
4. State or recognize the form of the data layout used for the Stratified Cox (SC) approaches for analyzing correlated data.
5. Given recurrent event data, outline the steps needed to analyze such data using the SC approaches.

Presentation

I. Overview



In this chapter we consider outcome events that may occur more than once over the follow-up time for a given subject. Such events are called "recurrent events." We focus on the **Counting Process (CP)** approach for analysis of such data that uses the Cox PH model, but we also describe alternative approaches that use a Stratified Cox (SC) PH model and a frailty model.

II. Examples of Recurrent Event Data

1. Multiple relapses from remission – leukemia patients
2. Repeated heart attacks – coronary patients
3. Recurrence of tumors – bladder cancer patients
4. Deteriorating episodes of visual acuity – macular degeneration patients

Up to this point, we have assumed that the event of interest can occur only once for a given subject. However, in many research scenarios in which the event of interest is not death, a subject may experience an event several times over follow-up. Examples of recurrent event data include:

1. Multiple episodes of relapses from remission comparing different treatments for leukemia patients.
2. Recurrent heart attacks of coronary patients being treated for heart disease.
3. Recurrence of bladder cancer tumors in a cohort of patients randomized to one of two treatment groups.
4. Multiple events of deteriorating visual acuity in patients with baseline macular degeneration, where each recurrent event is considered a more clinically advanced stage of a previous event.

Objective

Assess relationship of predictors to rate of occurrence, allowing for multiple events per subject

For each of the above examples, the event of interest differs, but nevertheless may occur more than once per subject. A logical objective for such data is to assess the relationship of relevant predictors to the rate in which events are occurring, allowing for multiple events per subject.

LEUKEMIA EXAMPLE

Do treatment groups differ in rates of relapse from remission?

In the leukemia example above, we might ask whether persons in one treatment group are experiencing relapse episodes at a higher rate than persons in a different treatment group.

HEART ATTACK EXAMPLE

Do smokers have a higher heart attack rate than nonsmokers?

If the recurrent event is a heart attack, we might ask, for example, whether smokers are experiencing heart attack episodes at a higher rate than nonsmokers.

LEUKEMIA AND HEART ATTACK EXAMPLES

All events are of the same type
The order of events is not important
Heart attacks: Treat as identical events;
Don't distinguish among 1st, 2nd, 3rd, etc. attack

For either of the above two examples, we are treating all events as if they were the same type. That is, the occurrence of an event on a given subject identifies the same disease without considering more specific qualifiers such as severity or stage of disease. We also are not taking into account the order in which the events occurred.

For example, we may wish to treat all heart attacks, whether on the same or different subjects, as identical types of events, and we don't wish to identify whether a given heart attack episode was the first, or the second, or the third event that occurred on a given subject.

BLADDER CANCER EXAMPLE

Compare overall tumor recurrence rate without considering order or type of tumor

The third example, which considers recurrence of bladder cancer tumors, can be considered similarly. That is, we may be interested in assessing the "overall" tumor recurrence rate without distinguishing either the order or type of tumor.

MACULAR DEGENERATION OF VISUAL ACUITY EXAMPLE

A second or higher event is more severe than its preceding event

The fourth example, dealing with macular degeneration events, however, differs from the other examples. The recurrent events on the same subject differ in that a second or higher event indicates a more severe degenerative condition than its preceding event.

Order of event is important

Consequently, the investigator in this scenario may wish to do separate analyses for each ordered event in addition to or instead of treating all recurrent events as identical.

Use a different analysis depending on whether

a. recurrent events are treated as identical
b. recurrent events involve different disease categories and/or the order of events is important

We have thus made an important distinction to be considered in the analysis of recurrent event data. If all recurrent events on the same subject are treated as identical, then the analysis required of such data is different than what is required if either recurrent events involve different disease categories and/or the order that events reoccur is considered important.

Recurrent events identical
⇓
Counting Process Approach
(Andersen et al., 1993)

The approach to analysis typically used when recurrent events are treated as identical is called the **Counting Process Approach** (Andersen et al., 1993).

Recurrent events: different disease categories or event order important
⇓
Stratified Cox (SC) Model Approaches

When recurrent events involve different disease categories and/or the order of events is considered important, a number of alternative approaches to analysis have been proposed that involve using stratified Cox (SC) models.

In this chapter, we focus on the **Counting Process (CP)** approach, but also describe the other stratified Cox approaches (in a later section).

III. Counting Process Example

Table 8.1. 2 Hypothetical Subjects Bladder Cancer Tumor Events

	Time interval	Event indicator	Treatment group
Al	0 to 3	1	1
	3 to 9	1	1
	9 to 21	1	1
	21 to 23	0	1
Hal	0 to 3	1	0
	3 to 15	1	0
	15 to 25	1	0

To illustrate the counting process approach, we consider data on two hypothetical subjects (Table 8.1), Al and Hal, from a randomized trial that compares two treatments for bladder cancer tumors.

Al gets recurrent bladder cancer tumors at months 3,9, and 21, and is without a bladder cancer tumor at month 23, after which he is no longer followed. Al received the treatment coded as 1.

Hal gets recurrent bladder cancer tumors at months 3, 15, and 25, after which he is no longer followed. Hal received the treatment coded as 0.

	Al	Hal
No. recurrent events	3	3
Follow-up time	23 months	25 months
Event times from start of follow-up	3, 9, 21	3, 15, 25
Additional months of follow-up after last event	2 months	0 months

Al has experienced 3 events of the same type (i.e., recurrent bladder tumors) over a follow-up period of 23 months. Hal has also experienced 3 events of the same type over a follow-up period of 25 months.

The three events experienced by Al occurred at different survival times (from the start of initial follow-up) from the three events experienced by Hal.

Also, Al had an additional 2 months of follow-up after his last recurrent event during which time no additional event occurred. In contrast, Hal had no additional event-free follow-up time after his last recurrent event.

Table 8.2. Example of Data Layout for Counting Process Approach

Subj	Interval Number	Time Start	Time Stop	Event Status	Treatment Group
Al	1	0	3	1	1
Al	2	3	9	1	1
Al	3	9	21	1	1
Al	4	21	23	0	1
Hal	1	0	3	1	0
Hal	2	3	15	1	0
Hal	3	15	25	1	0

In Table 8.2, we show for these 2 subjects, how the data would be set up for computer analyses using the counting process approach. Each subject contributes a line of data for each time interval corresponding to each recurrent event and any additional event-free follow-up interval. We previously introduced this format as the Counting Process (CP) data layout in section VI of Chapter 1.

Counting process: **Start** and **Stop** times

Standard layout: only **Stop (survival)** times (no recurrent events)

A distinguishing feature of the data layout for the counting process approach is that each line of data for a given subject lists the **start time** and **stop time** for each interval of follow-up. This contrasts with the standard layout for data with no recurrent events, which lists only the stop (**survival**) time.

Subj	Interval Number	Time Start	Time Stop	Event Status	Treatment Group
Sal	1	0	17	1	0
Mal	1	0	12	0	1

Note that if a third subject, Sal, failed without further events or follow-up occurring, then Sal contributes only one line of data, as shown at the left. Similarly, only one line of data is contributed by a (fourth) subject, Mal, who was censored without having failed at any time during follow-up.

IV. General Data Layout: Counting Process Approach

- N subjects
- r_i time intervals for subject i
- d_{ij} event staus (0 or 1) for subject i in interval j
- t_{ij0} start time for subject i in interval j
- t_{ij1} stop time for subject i in interval j
- X_{ijk} value of kth predictor for subject i in interval j
- $i = 1, 2, \ldots, N;\ j = 1, 2, \ldots, n_i;\ k = 1, 2, \ldots, p$

The general data layout for the counting process approach is presented in Table 8.3 for a dataset involving N subjects.

The ith subject has r_i recurrent events. d_{ij} denotes the event status (1 = failure, 0 = censored) for the ith subject in the jth time interval. t_{ij0} and t_{ij1} denote the start and stop times, respectively, for the ith subject in the jth interval. X_{ijk} denotes the value of the kth predictor for the ith subject in the jth interval.

Table 8.3. General Data Layout: CP Approach

Subject i	Interval j	status d_{ij}	Start t_{ij0}	Stop t_{ij1}	Predictors $X_{ij1} \ldots X_{ijp}$
1	1	d_{11}	t_{110}	t_{111}	$X_{111} \ldots X_{11p}$
1	2	d_{12}	t_{120}	t_{121}	$X_{121} \ldots X_{12p}$
.	.	.	.	.	. .
.	.	.	.	.	. .
.	.	.	.	.	. .
1	r_1	d_{1r_1}	t_{1r_10}	t_{1r_11}	$X_{1r_11} \ldots X_{1r_1p}$
.	.	.	.	.	. .
i	1	d_{i1}	t_{i10}	t_{i11}	$X_{i11} \ldots X_{i1p}$
i	2	d_{i2}	t_{i20}	t_{i21}	$X_{i21} \ldots X_{i2p}$
.	.	.	.	.	. .
.	.	.	.	.	. .
.	.	.	.	.	. .
i	r_i	d_{ir_i}	t_{ir_i0}	t_{ir_i1}	$X_{ir_i1} \ldots X_{ir_ip}$
.	.	.	.	.	. .
N	1	d_{N1}	t_{N10}	t_{N11}	$X_{N11} \ldots X_{N1p}$
N	2	d_{N2}	t_{N20}	t_{N21}	$X_{N21} \ldots X_{N2p}$
.	.	.	.	.	. .
.	.	.	.	.	. .
.	.	.	.	.	. .
N	r_N	d_{Nr_N}	t_{Nr_N0}	t_{Nr_N1}	$X_{Nr_N1} \ldots X_{Nr_Np}$

Subjects are not restricted to have the same number of time intervals (e.g., r_1 does not have to equal r_2) or the same number of recurrent events. If the last time interval for a given subject ends in censorship ($d_{ij} = 0$), then the number of recurrent events for this subject is $r_i - 1$; previous time intervals, however, usually end with a failure ($d_{ij} = 1$).

Also, start and stop times may be different for different subjects. (See the previous section's example involving two subjects.)

As with any survival data, the covariates (i.e., Xs) may be time-independent or time-dependent for a given subject. For example, if one of the Xs is "gender" (1 = female, 0 = male), the values of this variable will be all 1s or all 0s over all time intervals observed for a given subject. If another X variable is, say, a measure of daily stress level, the values of this variable are likely to vary over the time intervals for a given subject.

The second column ("Interval j") in the data layout is not needed for the **CP** analysis, but is required for other approaches described later.

Table 8.4 First 26 Subjects: Bladder Cancer Study

id	int	event	start	stop	tx	num	size
1	1	0	0	0	0	1	1
2	1	0	0	1	0	1	3
3	1	0	0	4	0	2	1
4	1	0	0	7	0	1	1
5	1	0	0	10	0	5	1
6	1	1	0	6	0	4	1
6	2	0	6	10	0	4	1
7	1	0	0	14	0	1	1
8	1	0	0	18	0	1	1
9	1	1	0	5	0	1	3
9	2	0	5	18	0	1	3
10	1	1	0	12	0	1	1
10	2	1	12	16	0	1	1
10	3	0	16	18	0	1	1
11	1	0	0	23	0	3	3
12	1	1	0	10	0	1	3
12	2	1	10	15	0	1	3
12	3	0	15	23	0	1	3
13	1	1	0	3	0	1	1
13	2	1	3	16	0	1	1
13	3	1	16	23	0	1	1
14	1	1	0	3	0	3	1
14	2	1	3	9	0	3	1
14	3	1	9	21	0	3	1
14	4	0	21	23	0	3	1
15	1	1	0	7	0	2	3
15	2	1	7	10	0	2	3
15	3	1	10	16	0	2	3
15	4	1	16	24	0	2	3
16	1	1	0	3	0	1	1
16	2	1	3	15	0	1	1
16	3	1	15	25	0	1	1
17	1	0	0	26	0	1	2
18	1	1	0	1	0	8	1
18	2	0	1	26	0	8	1
19	1	1	0	2	0	1	4
19	2	1	2	26	0	1	4
20	1	1	0	25	0	1	2
20	2	0	25	28	0	1	2
21	1	0	0	29	0	1	4
22	1	0	0	29	0	1	2
23	1	0	0	29	0	4	1
24	1	1	0	28	0	1	6
24	2	1	28	30	0	1	6
25	1	1	0	2	0	1	5
25	2	1	2	17	0	1	5
25	3	1	17	22	0	1	5
25	4	0	22	30	0	1	5
26	1	1	0	3	0	2	1
26	2	1	3	6	0	2	1
26	3	1	6	8	0	2	1
26	4	1	8	12	0	2	1
26	5	0	12	30	0	2	1

To illustrate the above general data layout, we present in Table 8.4 the data for the first 26 subjects from a study of recurrent bladder cancer tumors (Byar, 1980 and Wei, Lin, and Weissfeld, 1989). The entire dataset contained 86 patients, each followed for a variable amount of time up to 64 months.

The repeated event being analyzed is the recurrence of bladder cancer tumors after transurethral surgical excision. Each recurrence of new tumors was treated by removal at each examination.

About 25% of the 86 subjects experienced four events.

The exposure variable of interest is drug treatment status (**tx**, 0 = placebo, 1 = treatment with thiotepa). The covariates listed here are initial number of tumors (**num**) and initial size of tumors (**size**) in centimeters. The paper by Wei, Lin, and Weissfeld actually focuses on a different method of analysis (called "marginal"), which requires a different data layout than shown here. We later describe the "marginal" approach and its corresponding layout.

In these data, it can be seen that 16 of these subjects (id #s 1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 17, 18, 20, 21, 22, 23) had no recurrent events, 4 subjects had 2 recurrent events (id #s 10, 12, 19, 24), 4 subjects (id #s 13, 14, 16, 25) had 3 recurrent events, and 2 subjects (id #s 15, 26) had 4 recurrent events.

Moreover, 9 subjects (id #s 6, 9, 10, 12, 14, 18, 20, 25, 26) were observed for an additional event-free time interval after their last event. Of these, 4 subjects (id #s 6, 9, 18, 20) experienced only one event (i.e., no recurrent events).

V. The Counting Process Model and Method

Cox PH Model

$h(t, \mathbf{X}) = h_0(t)\exp[\Sigma\beta_i X_i]$

Need to

Assess PH assumption for X_i
Consider stratified Cox or extended Cox if PH assumption not satisfied
Use extended Cox for time-dependent variables

The model typically used to carry out the **Counting Process** approach is the standard Cox PH model, once again shown here at the left.

As usual, the PH assumption needs to be evaluated for any time-independent variable. A stratified Cox model or an extended Cox model would need to be used if one or more time-independent variables did not satisfy the PH assumption. Also, an extended Cox model would be required if inherently time-dependent variables were considered.

The primary difference in the way the Cox model is used for analyzing recurrent event data versus nonrecurrent (one time interval per subject) data is the way several time intervals on the same subject are treated in the formation of the likelihood function maximized for the Cox model used.

Recurrent event data	Nonrecurrent event data
(Likelihood function formed differently)	
Subjects with > 1 time interval remain in the risk set until last interval is completed	Subjects removed from risk set at time of failure or censorship
Different lines of data are treated as **independent** even though several outcomes on the **same** subject	Different lines of data are treated as **independent** because they come from **different** subjects

To keep things simple, we assume that the data involve only time-independent variables satisfying the PH assumption. For recurrent survival data, a subject with more than one time interval remains in the risk set until his or her last interval, after which the subject is removed from the risk set. In contrast, for nonrecurrent event data, each subject is removed from the risk set at the time of failure or censorship.

Nevertheless, for subjects with two or more intervals, the different lines of data contributed by the same subject are treated in the analysis as if they were independent contributions from different subjects, even though there are several outcomes on the **same** subject.

In contrast, for the standard Cox PH model approach for nonrecurrent survival data, different lines of data are treated as independent because they come from **different** subjects.

Cox PH Model for CP Approach: Bladder Cancer Study

For the bladder cancer study described in Table 8.4, the basic Cox PH model fit to these data takes the form shown at the left.

$$h(t, \mathbf{X}) = h_0(t)\mathbf{exp}[\beta\ \mathbf{tx} + \gamma_1\ \mathbf{num} + \gamma_2\ \mathbf{size}]$$

where

tx = 1 if thiotepa, 0 if placebo
num = initial # of tumors
size = initial size of tumors

The primary (exposure) variable of interest in this model is the treatment variable **tx**. The variables **num** and **size** are considered as potential con-founders. All three variables are time-independent variables.

No-interaction Model

Interaction model would involve product terms

tx × **num** and/or **tx** × **size**

This is a no-interaction model because it does not contain product terms of the form **tx** × **num** or **tx** × **size**. An interaction model involving such product terms could also be considered, but we only present the no-interaction model for illustrative purposes.

Table 8.5. Ordered Failure Time and Risk Set Information for First 26 Subjects in Bladder Cancer Study

Ordered failure times $t_{(f)}$	# in risk set n_f	# failed m_f	# censored in $[t_{(f)}, t_{(f+1)})$	Subject ID #s for outcomes in $[t_{(f)}, t_{(f+1)})$
0	**26**	–	1	1
1	25	1	1	2, 18
2	24	2	0	19, 25
3	24	4	1	3, 13, 14, 16, 26
5	23	1	0	9
6	23	2	0	6, 26
7	23	1	1	4, 15
8	22	1	0	26
9	22	1	0	14
10	22	2	2	5, 6, 12, 15
12	20	2	1	7, 10, 26
15	19	2	0	12, 16
16	19	3	0	10, 13, 15
17	19	1	3	8, 9, 10, 25
21	16	1	0	14
22	16	1	0	25
23	16	1	3	11, 12, 13, 14
24	12	1	0	15
25	11	2	0	16, 20
26	10	1	2	17, 18, 19
28	7	1	4	20, 21, 22, 23, 24
30	3	1	2	24, 25, 26
		32	21	

Table 8.5 at the left provides ordered failure times and corresponding risk set information that would result if the first 26 subjects that we described in Table 8.4 comprised the entire dataset. (Recall that there are 86 subjects in the complete dataset.)

Because we consider 26 subjects, the number in the risk set at ordered failure time $t_{(0)}$ is $n_0 = 26$. As these subjects fail (i.e., develop a bladder cancer tumor) or are censored over follow-up, the number in the risk set will decrease from the fth to the f + 1th ordered failure time provided that no subject who fails at time $t_{(f)}$ either has a recurrent event at a later time or has additional follow-up time until later censorship. In other words, *a subject who has additional follow-up time after having failed at $t_{(f)}$ does not drop out of the risk set after $t_{(f)}$.*

Table 8.6. Focus on Subject #s **19** and **25** from Table 8.5

$t_{(f)}$	$n_{(f)}$	$m_{(f)}$	$q_{(f)}$	Subject ID #s
0	26	–	1	1
1	25	1	1	2, 18
2	**24**	2	0	**19, 25**
3	**24**	4	1	3, 13, 14, 16, 26
			•	
			•	
			•	
17	19	1	3	8, 9, 10, **25**
21	16	1	0	14
22	16	1	0	**25**
23	16	1	3	11, 12, 13, 14
24	12	1	0	15
25	11	2	0	16, 20
26	10	1	2	17, 18, **19**
28	7	1	4	20, 21, 22, 23, 24
30	3	1	2	24, **25**, 26

For example, at month $t_{(f)} = 2$, subject #s **19** and **25** fail, but the number in the risk set at that time ($n_f = 24$) does not decrease (by 2) going into the next failure time because each of these subjects has later recurrent events. In particular, subject #**19** has a recurrent event at month $t_{(f)} = 26$ and subject #**25** has two recurrent events at months $t_{(f)} = 17$ and $t_{(f)} = 22$ and has additional follow-up time until censored month 30.

Table 8.7. Focus on Subject #s **3, 13, 14, 16, 26** from Table 8.5

$t_{(f)}$	$n_{(f)}$	$m_{(f)}$	$q_{(f)}$	Subject ID #s
0	26	–	1	1
1	25	1	1	2, 18
2	24	2	0	19, 25
3	**24**	4	1	**3, 13, 14, 16, 26**
5	**23**	1	0	9
6	23	2	0	6, **26**
7	23	1	1	4, 15
8	22	1	0	**26**
9	22	1	0	**14**
10	22	2	2	5, 6, 12, 15
12	20	2	1	7, 10, **26**
15	19	2	0	12, **16**
16	19	3	0	10, **13**, 15
17	19	1	3	8, 9, 10, 25
21	16	1	0	14
22	16	1	0	25
23	16	1	3	11, 12, **13, 14**
24	12	1	0	15
25	11	2	0	**16**, 20
26	10	1	2	17, 18, 19
28	7	1	4	20, 21, 22, 23, 24
30	3	1	2	24, 25, **26**

As another example from Table 8.5, subject #s **3, 13, 14, 16, 26** contribute information at ordered failure time $t_{(f)} = 3$, but the number in the risk set only drops from 24 to 23 even though the last four of these subjects all fail at $t_{(f)} = 3$. Subject #**3** is censored at month 4 (see Table 8.4), so this subject is removed from the risk set after failure time $t_{(f)} = 3$. However, subjects **13, 14, 16,** and **26** all have recurrent events after $t_{(f)} = 3$, so they are not removed from the risk set after $t_{(f)} = 3$.

Subject #26 appears in the last column 5 times. This subject contributes 5 (start, stop) time intervals, fails at months 3, 6, 8, and 12, and is also followed until month 30, when he is censored.

"Gaps" in follow-up time:

0	10 lost	**gap**	25 re-enter	50

Another situation, which is not illustrated in these data, involves "gaps" in a subject's follow-up time. A subject may leave the risk set (e.g., lost to follow-up) at, say, time = 10 and then re-enter the risk set again and be followed from, say, time = 25 to time = 50. This subject has a follow-up gap during the period from time = 10 to time = 25.

No Interaction Cox PH Model

$$h(t,\mathbf{X}) = h_0(t)\mathbf{exp}[\beta\ \mathbf{tx} + \gamma_1\ \mathbf{num} + \gamma_2\ \mathbf{size}]$$

Partial likelihood function:

$$\mathbf{L} = \mathbf{L_1} \times \mathbf{L_2} \times \cdots \times \mathbf{L_{22}}$$

The (partial) likelihood function (**L**) used to fit the no-interaction Cox PH model is expressed in typical fashion as the product of individual likelihoods contributed by each ordered failure time and corresponding risk set information in Table 8.5. There are 22 such terms in this product because there are 22 ordered failure times listed in Table 8.5.

$\mathbf{L_f}$ = individual likelihood at $t_{(j)}$
= Pr[failing at $t_{(f)}$ | survival up to $t_{(f)}$]
$f = 1, 2, \ldots, 22$

Each individual likelihood L_f essentially gives the conditional probability of failing at time $t_{(f)}$, given survival (i.e., remaining in the risk set) at $t_{(f)}$.

$$\mathbf{L_f} = \frac{\mathbf{exp}\left(\beta tx_{(f)} + \gamma_1 num_{(f)} + \gamma_2 size_{(f)}\right)}{\sum\limits_{s \text{ in } R\left(t_{(f)}\right)} \mathbf{exp}\left(\beta tx_{s(f)} + \gamma_1 num_{s(f)} + \gamma_2 size_{s(f)}\right)}$$

If there is only one failure at the jth ordered failure time, L_f is expressed as shown at the left for the above no-interaction model. In this expression $tx_{(f)}$, $num_{(f)}$, and $size_{(f)}$ denote the values of the variables **tx, num,** and **size** for the subject failing at month $t_{(f)}$.

$tx_{(f)}$, $num_{(f)}$, and $size_{(f)}$ values of **tx, num**, and **size** at $t_{(f)}$

$tx_{s(f)}$, $num_{s(f)}$, and $size_{s(f)}$ values of **tx, num**, and **size** for subject s in $R(t_{(f)})$

The terms $tx_{s(f)}$, $num_{s(f)}$, and $size_{s(f)}$ denote the values of the variables **tx, num,** and **size** for the subject s in the risk set $R(t_{(f)})$. Recall that $R(t_{(f)})$ consists of all subjects remaining at risk at failure time $t_{(f)}$.

Data for Subject #25

id	int	event	start	stop	tx	num	size
25	1	1	0	2	0	1	5
25	2	1	2	17	0	1	5
25	3	1	17	**22**	0	1	5
25	4	0	22	30	0	1	5

f = 15th ordered failure time
n_{15} = 16 subjects in risk set at $t_{(15)} = 22$:

$R(t_{(15)} = 22)$ = {subject #s 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26}

For example, subject #25 from Table 8.4 failed for the third time at month 22, which is the f = 15th ordered failure time listed in Table 8.5. It can be seen that $n_f = 16$ of the initial 26 subjects are still at risk at the beginning of month 22. The risk set at this time includes subject #25 and several other subjects (#s 12, 13, 14, 15, 16, 18, 19, 26) who already had at least one failure prior to month 22.

$$\mathbf{L}_{15} = \frac{\mathbf{exp}(\beta(0) + \gamma_1(1) + \gamma_2(5))}{\sum_{s \text{ in } R(t_{(15)})} \mathbf{exp}(\beta tx_{s(15)} + \gamma_1 num_{s(15)} + \gamma_1 size_{s(15)})}$$

The corresponding likelihood $\mathbf{L}_{15}$ at $t_{(15)} = 22$ is shown at the left. Subject #25's values $tx_{25(15)} = 0$, $num_{25(15)} = 1$, and $size_{25(15)} = 5$, have been inserted into the numerator of the formula. The denominator will contain a sum of 16 terms, one for each subject in the risk set at $t_{(15)} = 22$.

Computer program formulates partial likelihood **L**
(See Computer Appendix)

The overall partial likelihood **L** will be formulated internally by the computer program once the data layout is in the correct form and the program code used involves the (start, stop) formulation.

VI. Robust Estimation

Data for Subject #14

id	int	event	start	stop	tx	num	size
14	1	1	0	3	0	3	1
14	2	1	3	9	0	3	1
14	3	1	9	21	0	3	1
14	4	0	21	23	0	3	1

As illustrated for subject #14 at the left, each subject contributes a line of data for each time interval corresponding to each recurrent event and any additional event-free follow-up interval.

Up to this point:
the 4 lines of data for subject #14 are treated as independent observations

We have also pointed out that the Cox model analysis described up to this point treats different lines of data contributed by the same subject as if they were independent contributions from different subjects.

Nevertheless,

- Observations of the same subject are correlated
- Makes sense to adjust for such correlation in the analysis

Nevertheless, it makes sense to view the different intervals contributed by a given subject as representing correlated observations on the same subject that must be accounted for in the analysis.

Robust (Empirical) Estimation

- Adjusts
  $\widehat{\mathbf{Var}}(\hat{\beta}_k)$
  where
  $\hat{\beta}_k$
  is an estimated regression coefficient
- accounts for misspecification of assumed correlation structure

A widely used technique for adjusting for the correlation among outcomes on the same subject is called **robust estimation** (also referred to as **empirical estimation**). This technique essentially involves adjusting the estimated variances of regression coefficients obtained for a fitted model to account for misspecification of the correlation structure assumed (see Zeger and Liang, 1986 and Kleinbaum and Klein, 2010).

CP approach: assumes independence

Goal of robust estimation: adjust for correlation within subjects

In the CP approach, the assumed correlation structure is independence; that is, the Cox PH model that is fit assumes that different outcomes on the same subject are independent. Therefore the goal of robust estimation for the CP approach is to obtain variance estimators that adjust for correlation within subjects when previously no such correlation was assumed.

Same goal for other approaches for analyzing recurrent event data

This is the same goal for other approaches for analyzing recurrent event data that we describe later in this chapter.

Do not adjust

$\hat{\beta}_k$

Only adjust

$\widehat{\mathbf{Var}}\left(\hat{\beta}_k\right)$

Note that the estimated regression coefficients themselves are not adjusted; only the estimated variances of these coefficients are adjusted.

Robust (Empirical) Variance

allows
tests of hypotheses and confidence intervals
that account for correlated data

The robust (i.e., empirical) estimator of the variance of an estimated regression coefficient therefore allows tests of hypotheses and confidence intervals about model parameters that account for correlation within subjects.

Matrix formula:

derived from ML estimation

We briefly describe the formula for the robust variance estimator below. This formula is in matrix form and involves terms that derive from the set of "score" equations that are used to solve for ML estimates of the regression coefficients. This information may be of interest to the more mathematically inclined reader with some background in methods for the analysis of correlated data (Kleinbaum and Klein, 2010).

Formula not essential for using computer packages

However, the information below is not essential for an understanding of how to obtain robust estimators using computer packages. (See Computer Appendix.)

Extension (Lin and Wei, 1989) of **information sandwich estimator** (Zeger and Liang, 1986)

The robust estimator for recurrent event data was derived by Lin and Wei (1989) as an extension similar to the "information sandwich estimator" proposed by Zeger and Liang (1986) for generalized linear models. SAS and Stata use variations of this estimator that give slightly different estimates.

Matrix formula

$$\hat{\mathbf{R}}\left(\hat{\beta}\right) = \widehat{\mathbf{Var}}\left(\hat{\beta}\right)\left[\hat{\mathbf{R}}_S'\hat{\mathbf{R}}_S\right]\widehat{\mathbf{Var}}\left(\hat{\beta}\right)$$

where

$\widehat{\mathbf{Var}}\left(\hat{\beta}\right)$

is the **information matrix**, and

$\hat{\mathbf{R}}_S$

is matrix of **score residuals**

The general form of this estimator can be most conveniently written in matrix notation as shown at the left. In this formula, the variance expression denotes the **information matrix** form of estimated variances and covariances obtained from (partial) ML estimation of the Cox model being fit. The $\hat{\mathbf{R}}_S$ expression in the middle of the formula denotes the matrix of **score residuals** obtained from ML estimation.

Formula applies to other approaches for analyzing recurrent event data

The robust estimation formula described above applies to the **CP** approach as well as other approaches for analyzing recurrent event data described later in this chapter.

VII. Results for CP Example

We now describe the results from using the **CP** approach on the Bladder Cancer Study data involving all 85 subjects.

Table 8.8. Edited SAS Output from CP Approach on Bladder Cancer Data (N = 85 Subjects) Without Robust Variances

Var	DF	Parameter Estimate	Std Error	Chisq	P	$\widehat{HR}$
tx	**1**	**−0.4071**	**0.2001**	**4.140**	**0.042**	**0.667**
num	1	0.1607	0.0480	11.198	0.001	1.174
size	1	−0.0401	0.0703	0.326	0.568	0.961

−2 LOG L = 920.159

Table 8.8 gives edited output from fitting the no-interaction Cox PH model involving the three predictors **tx, num,** and **size**. A likelihood ratio chunk test for interaction terms **tx** × **num** and **tx** × **size** was nonsignificant, thus supporting the no-interaction model shown here. The PH assumption was assumed satisfied for all three variables.

Table 8.9. Robust Covariance Matrix, CP Approach on Bladder Cancer Data

	tx	**num**	**size**
tx	**0.05848**	−0.00270	−0.00051
num	−0.00270	0.00324	0.00124
size	−0.00051	0.00124	0.00522

Table 8.9 provides the covariance matrix obtained from robust estimation of the variances of the estimated regression coefficients of **tx, num,** and **size**. The values on the diagonal of this matrix give robust estimates of these variances and the off-diagonal values give covariances.

Robust standard error for **tx**
= square-root (**.05848**) = **0.2418**

Non**robust standard error for tx**
= **0.2001**

Because the exposure variable of interest in this study is **tx,** the most important value in this matrix is **0.05848**. The square root of this value is **0.2418,** which gives the robust standard error of the estimated coefficient of the **tx** variable. Notice that this robust estimator is similar but somewhat different from the nonrobust estimator of **0.2001** shown in Table 8.8.

Summary of Results from CP Approach

Hazard Ratio **tx**: exp(−0.407) = 0.667 (= 1/1.5)

Wald Chi-Square **tx**:	robust	nonrobust
	2.83	4.14
P-value **tx**:	.09	.04

(H_0: no effect of **tx,** H_A: two sided)

We now summarize the **CP** results for the effect of the exposure variable **tx** on recurrent event survival controlling for **num** and **size**. The hazard ratio estimate of 0.667 indicates that the hazard for the placebo is 1.5 times the hazard for the treatment.

95% CI for HR **tx** (robust): (0.414, 1.069)

H_A: one-sided, both p-values < .05

Using robust estimation, the Wald statistic for this hazard ratio is borderline nonsignificant (P = .09). Using the nonrobust estimator, the Wald statistic is borderline significant (P = .04). Both these P-values, however, are for a two-sided alternative. For a one-sided alternative, both P-values would be significant at the .05 level. The 95% confidence interval using the robust variance estimator is quite wide in any case.

We return to the analysis of these data when we discuss other approaches for analysis of recurrent event data.

VIII. Other Approaches Stratified Cox

3 stratified Cox (SC) approaches:

Stratified CP (Prentice, Williams and
Gap Time Peterson, 1981)
Marginal (Wei, Lin, and Weissfeld, 1989)

Goal: distinguish order of recurrent events

We now describe three other approaches for analyzing recurrent event data, each of which uses a Stratified Cox (SC) PH model. They are called **Stratified CP, Gap Time,** and **Marginal**. These approaches are often used to distinguish the order in which recurrent events occur.

Strata variable: time interval # treated as categorical

The "strata" variable for each approach treats the time interval number as a categorical variable.

Example:
maximum of **4** failures per subject
⇓
Strata variable = 1 for time interval # 1
2 for time interval # 2
3 for time interval # 3
4 for time interval # 4

For example, if the maximum number of failures that occur on any given subject in the dataset is, say, 4, then time interval #1 is assigned to stratum 1, time interval #2 to stratum 2, and so on.

Time **between** two events:

Stratified CP $\frac{0 \quad 50 \rightarrow 80}{\text{entry}}$

Gap Time $\frac{0 \rightarrow 30}{\text{ev1} \quad \text{ev2}}$

Both **Stratified CP** and **Gap Time** approaches focus on survival time **between** two events. However, **Stratified CP** uses the actual times of the two events from study entry, whereas **Gap Time** starts survival time at 0 for the earlier event and stops at the later event.

Marginal

- Total survival time from study entry until kth event
- Recurrent events of different types

The **Marginal** approach, in contrast to each conditional approach, focuses on **total** survival time from study entry until the occurrence of a specific (e.g., kth) event; this approach is suggested when recurrent events are viewed to be of different types.

Stratified CP for Subject 10

id	int	event	start	stop	tx	num	size
10	1	1	0	12	0	1	1
10	2	1	12	16	0	1	1
10	3	0	16	18	0	1	1

The **stratified CP** approach uses the exact same (start, stop) data layout format used for the **CP** approach, except that for **Stratified CP,** an SC model is used rather than a standard (unstratified) PH model. The strata variable here is **int** in this listing.

Gap Time for Subject 10

(stop = Interval Length Since Previous Event)

id	int	event	start	stop	tx	num	size
10	1	1	0	12	0	1	1
10	2	1	0	4	0	1	1
10	3	0	0	2	0	1	1

The **Gap Time** approach also uses a (start, stop) data layout, but the start value is always 0 and the stop value is the time interval length since the previous event. The model here is also a SC model.

Marginal approach
Standard (nonrecurrent event) layout, i.e., without (start, stop) columns

The **Marginal** approach uses the standard (nonrecurrent event) data layout instead of the (start, stop) layout, as illustrated below.

Marginal Approach for Subject 10

id	int	event	stime	tx	num	size
10	1	1	12	0	1	1
10	2	1	16	0	1	1
10	3	0	18	0	1	1
10	4	0	18	0	1	1

The **Marginal** approach layout, shown at the left, contains four lines of data in contrast to the three lines of data that would appear for subject #10 using the **CP, Stratified CP,** and **Gap Time** approaches

Marginal approach
Each subject at risk for all failures that might occur

actual failures ≤ # possible failures

The reason why there are 4 lines of data here is that, for the **Marginal** approach, each subject is considered to be at risk for all failures that might occur, regardless of the number of events a subject actually experienced.

Bladder cancer data:

Maximum # (possible) failures = 4

So, subject 10 (as well as all other subjects) gets 4 lines of data

Because the maximum number of failures being considered in the bladder cancer data is 4 (e.g., for subject #s 15 and 26), subject #10, who failed only twice, will have two additional lines of data corresponding to the two additional failures that could have possibly occurred for this subject.

Fundamental Difference Among the 3 SC Approaches

Risk set differs for strata after first event

The three alternative SC approaches (**Stratified CP, Gap Time,** and **Marginal**) fundamentally differ in the way the risk set is determined for strata corresponding to events after the first event.

Gap Time: time until 1st event does not influence risk set for later events (i.e., clock reset to 0 after event occurs)

With **Gap Time,** the time until the first event does not influence the composition of the risk set for a second or later event. In other words, the clock for determining who is at risk gets reset to 0 after each event occurs.

Stratified CP: time until 1st event influences risk set for later events

In contrast, with **Stratified CP,** the time until the first event affects the composition of the risk set for later events.

Marginal: risk set determined from time since study entry

With the **Marginal** approach, the risk set for the kth event (k = 1, 2, ...) identifies those at risk for the kth event since entry into the study.

EXAMPLE

			Days		
ID	**Status**	**Stratum**	**Start**	**Stop**	**tx**
M	1	1	0	100	1
M	1	2	100	105	1
H	1	1	0	30	0
H	1	2	30	50	0
P	1	1	0	20	0
P	1	2	20	60	0
P	1	3	60	85	0

Suppose, for example, that Molly (M), Holly (H), and Polly (P) are the only three subjects in the dataset shown at the left. Molly receives the treatment (tx = 1) whereas Holly and Polly receive the placebo (tx = 0). All three have recurrent events at different times. Also, Polly experiences three events whereas Molly and Holly experience two.

Stratified CP

Stratum 1			Stratum 2		
$t_{(f)}$	n_f	$R(t_{(f)})$	$t_{(f)}$	n_f	$R(t_{(f)})$
0	3	{M, H, P}	20	1	{P}
20	3	{M, H, P}	30	2	{H, P}
30	2	{M, H}	50	2	{H, P}
100	1	{M}	60	1	{P}
			105	1	{M}

The table at the left shows how the risk set changes over time for strata 1 and 2 if the **Stratified CP** approach is used. For **stratum 2**, there are no subjects in the risk set until t = 20, when Polly gets the earliest first event and so becomes at risk for a second event. Holly enters the risk set at t = 30. So at t = 50, when the earliest second event occurs, the risk set contains Holly and Polly. Molly is not at risk for getting her second event until t = 100. The risk set at t = 60 contains only Polly because Holly has already had her second event at t = 50. And the risk set at t = 105 contains only Molly because both Holly and Polly have already had their second event by t = 105.

Gap Time

Stratum 1			Stratum 2		
$t_{(f)}$	n_f	$R(t_{(f)})$	$t_{(f)}$	n_f	$R(t_{(f)})$
0	3	{M, H, P}	0	3	{M, H, P}
20	3	{M, H, P}	5	3	{M, H, P}
30	2	{M, H}	20	2	{H, P}
100	1	{M}	40	1	{P}

The next table shows how the risk set changes over time if the **Gap Time** approach is used. Notice that the data for **stratum 1** are identical to those for **Stratified CP**. For **stratum 2,** however, all three subjects are at risk for the second event at t = 0 and at t = 5, when Molly gets her second event 5 days after the first occurs. The risk set at t = 20 contains Holly and Polly because Molly has already had her second event by t = 20. And the risk set at t = 40 contains only Polly because both Molly and Holly have already had their second event by t = 40.

Marginal

Stratum 1			Stratum 2		
$t_{(f)}$	n_f	$R(t_{(f)})$	$t_{(f)}$	n_f	$R(t_{(f)})$
0	3	{M, H, P}	0	3	{M, H, P}
20	3	{M, H, P}	50	3	{M, H, P}
30	2	{M, H}	60	2	{M, P}
100	3	{M}	105	1	{M}

We next consider the **Marginal** approach. For **stratum 1**, the data are identical again to those for **Stratified CP**. For **stratum 2,** however, all three subjects are at risk for the second event at t = 0 and at t = 50, when Holly gets her second event. The risk set at t = 60 contains Molly and Polly because Holly has already had her second event at t = 50. And the risk set at t = 105 contains only Molly because both Holly and Polly have already had their second event by t = 60.

Stratum 3 for Marginal approach follows

Because Polly experienced three events, there is also a third stratum for this example, which we describe for the marginal approach only.

Marginal
Stratum 3

$t_{(f)}$	n_f	$R(t_{(f)})$
0	3	{M, H, P}
85	2	{M, P}

Note: H censored by t = 85

Using the marginal approach, all three subjects are considered at risk for the third event when they enter the study (t = 0), even though Molly and Holly actually experience only two events. At t = 85, when Polly has her third event, Holly, whose follow-up ended at t = 50, is no longer in the risk set; which still includes Molly because Molly's follow-up continues until t = 105.

Basic idea (**Marginal** approach):

Each failure considered a separate process

Allows stratifying on

- Failure order
- Different failure type (e.g., stage 1 vs. stage 2 cancer)

The basic idea behind the **Marginal** approach is that it allows each failure to be considered as a separate process. Consequently, the **Marginal** approach not only allows the investigator to consider the ordering of failures as separate events (i.e., strata) of interest, but also allows the different failures to represent different types of events that may occur on the same subject.

Stratified Cox PH (SC) Model for all 3 alternative approaches

Use standard computer program for SC (e.g., SAS's PHREG, Stata's stcox, SPSS's coxreg, R's Coxph)

All three alternative approaches, although differing in the form of data layout and the way the risk set is determined, nevertheless use a stratified Cox PH model to carry out the analysis. This allows a standard program that fits a SC model (e.g., SAS's PHREG) to perform the analysis.

No-interaction SC model for bladder cancer data

$$h_g(t,\mathbf{X}) = h_{0g}(t)\mathbf{exp}[\beta\ \mathbf{tx} + \gamma_1\ \mathbf{num} + \gamma_2\ \mathbf{size}]$$

where g = 1, 2, 3, 4

The models used for the three alternative SC approaches are therefore of the same form. For example, we show on the left the no-interaction SC model appropriate for the bladder cancer data we have been illustrating.

Two types of SC models:

No-interaction versus interaction model
Typically compared using LR test

As described previously in Chapter 5 on the stratified Cox procedure, a no-interaction stratified Cox model is not appropriate if there is interaction between the stratified variables and the predictor variables put into the model. Thus, it is necessary to assess whether an interaction version of the SC model is more appropriate, as typically carried out using a likelihood ratio test.

Version 1: Interaction SC Model

$$h_g(t,\mathbf{X}) = h_{0g}(t)\,\mathbf{exp}[\beta_g\,\mathbf{tx} + \gamma_{1g}\,\mathbf{num} + \gamma_{2g}\,\mathbf{size}]$$

$g = 1, 2, 3, 4$

For the bladder cancer data, we show at the left two equivalent versions of the SC interaction model. The first version separates the data into 4 separate models, one for each stratum.

Version 2: Interaction SC Model

$$\begin{aligned} h_g(t,\mathbf{X}) = h_{0g}(t)\mathbf{exp}[&\beta\,\mathbf{tx} + \gamma_1\,\mathbf{num} \\ &+ \gamma_2\,\mathbf{size} + \delta_{11}(Z_1^* \times \mathbf{tx}) \\ &+ \delta_{12}(Z_2^* \times \mathbf{tx}) + \delta_{13}(Z_3^* \times \mathbf{tx}) \\ &+ \delta_{21}(Z_1^* \times \mathbf{num}) + \delta_{22}(Z_2^* \times \mathbf{num}) \\ &+ \delta_{23}(Z_3^* \times \mathbf{num}) + \delta_{31}(Z_1^* \times \mathbf{size}) \\ &+ \delta_{32}(Z_2^* \times \mathbf{size}) + \delta_{33}(Z_3^* \times \mathbf{size})] \end{aligned}$$

where Z_1^*, Z_2^*, and Z_3^* are 3 dummy variables for the 4 strata.

The second version contains product terms involving the stratified variable with each of the 3 predictors in the model. Because there are 4 strata, the stratified variable is defined using 3 dummy variables Z_1^*, Z_2^*, and Z_3^*.

H_0 (Version 1)

$$\begin{aligned} \beta_1 &= \beta_2 = \beta_3 = \beta_4 \equiv \beta, \\ \gamma_{11} &= \gamma_{12} = \gamma_{13} = \gamma_{14} \equiv \gamma_1, \\ \gamma_{21} &= \gamma_{22} = \gamma_{23} = \gamma_{24} \equiv \gamma_2 \end{aligned}$$

H_0 (Version 2)

$$\begin{aligned} \delta_{11} &= \delta_{12} = \delta_{13} = \delta_{21} = \delta_{22} \\ &= \delta_{23} = \delta_{31} = \delta_{32} = \delta_{33} \\ &= 0 \end{aligned}$$

The null hypotheses for the LR test that compares the interaction with the no-interaction SC model is shown at the left for each version. The df for the LR test is 9.

Interaction SC model may be used regardless of LR test result

- Allows separate HRs for **tx** for each stratum
- if no-interaction SC, then only an overall effect of **tx** can be estimated

Even if the no-interaction SC model is found more appropriate from the likelihood ratio test, the investigator may still wish to use the interaction SC model in order to obtain and evaluate different hazard ratios for each stratum. In other words, if the no-interaction model is used, it is not possible to separate out the effects of predictors (e.g., **tx**) within each stratum, and only an overall effect of a predictor on survival can be estimated.

Recommend using

robust estimation

$$\hat{\mathbf{R}}\left(\hat{\beta}\right) = \widehat{\mathbf{Var}}\left(\hat{\beta}\right)\left[\hat{\mathbf{R}}'_S\hat{\mathbf{R}}_S\right]\widehat{\mathbf{Var}}\left(\hat{\beta}\right)$$

to adjust for correlation of observations on the same subject

For each of the SC alternative approaches, as for the **CP** approach, it is recommended to use **robust estimation** to adjust the variances of the estimated regression coefficients for the correlation of observations on the same subject. The general form for the robust estimator is the same as in the **CP** approach, but will give different numerical results because of the different data layouts used in each method.

IX. Bladder Cancer Study Example (Continued)

We now present and compare SAS results from using all four methods we have described – **CP, Stratified CP, Gap Time,** and **Marginal** – for analyzing the recurrent event data from the bladder cancer study.

Table 8.10. Estimated βs and HRs for **tx** from Bladder Cancer Data

Model	$\hat{\beta}$	$\widehat{HR} = \exp(\hat{\beta})$
CP	−0.407	0.666 (=1/1.50)
SCP	−0.334	0.716 (=1/1.40)
GT	−0.270	0.763 (=1/1.31)
M	−0.580	0.560 (=1/1.79)

CP = Counting Process,
SCP = Stratified CP
GT = Gap Time, M = Marginal

Table 8.10 gives the regression coefficients for the **tx** variable and their corresponding hazard ratios (i.e., **exp**($\hat{\beta}$) for the no-interaction Cox PH models using these four approaches). The model used for the **CP** approach is a standard Cox PH model whereas the other three models are SC models that stratify on the event order.

HR for **M: 0.560** (=1/1.79)
differs from
HRs for **CP: 0.666** (=1/1.50),
SCP: 0.716 (=1/1.40),
GT: 0.763 (=1/1.31)

From this table, we can see that the hazard ratio for the effect of the exposure variable **tx** differs somewhat for each of the four approaches, with the **Marginal** model giving a much different result from that obtained from the other three approaches.

Table 8.11 Estimated βs, SE(β)s, and P-Values for **tx** from No-Interaction Model for Bladder Cancer Data

Model	$\hat{\beta}$	SE(NR)	SE(R)	P(NR)	P(R)
CP	−0.407	0.200	0.242	.042	.092
SCP	−0.334	0.216	0.197	.122	.090
GT	−0.270	0.208	00.208	.195	.194
M	−0.580	0.201	0.303	.004	.056

CP = Counting Process, SCP = Stratified CP,
GT = Gap Time, M = Marginal,
NR = Nonrobust, R = Robust, P = Wald P-value

Table 8.11 provides, again for the exposure variable **tx** only, the regression coefficients, robust standard errors, nonrobust standard errors, and corresponding Wald Statistic P-values obtained from using the no-interaction model with each approach.

SE(NR) differs from **SE(R)**
P(NR) differs from **P(R)**
but no clear pattern

for example,
CP: P(NR) = .042 < P(R) = .092
SCP: P(NR) = .122 > P(R) = .090
GT: P(NR) = .195 = P(R) = .194

The nonrobust and robust standard errors and P-values differ to some extent for each of the different approaches. There is also no clear pattern to suggest that the nonrobust results will always be either higher or lower than the corresponding robust results.

Wald test statistic(s):

$$Z = \hat{\beta}/SE(\hat{\beta}) \Leftrightarrow Z^2 = [\hat{\beta}/SE(\hat{\beta})]^2$$
$$\sim N(0,1) \textbf{ under } \mathbf{H_0}\textbf{: } \boldsymbol{\beta} = \mathbf{0} \sim \chi^2_{1\ df}$$

The P-values shown in Table 8.11 are computed using the standard Wald test Z or chi-square statistic, the latter having a chi-square distribution with 1 df under the null hypothesis that there is no effect of **tx**.

Table 8.12 Estimated βs and Robust SE(β)s for tx from Interaction SC Model for Bladder Cancer Data

	Interaction SC Model				
Model	Str1 $\hat{\beta}_1$ (SE)	Str2 $\hat{\beta}_2$ (SE)	Str3 $\hat{\beta}_3$ (SE)	Str4 $\hat{\beta}_4$ (SE)	No Interaction $\hat{\beta}$ (SE)
CP	—	—	—	—	−.407 (.242)
SCP	−.518 (.308)	−.459 (.441)	.117 (.466)	−.041 (.515)	−.334 (.197)
GT	−.518 (.308)	−.259 (.402)	.221 (.620)	−.195 (.628)	−.270 (.208)
M	−.518 (.308)	−.619 (.364)	−.700 (.415)	−.651 (.490)	−.580 (.303)

CP = Counting Process, SCP = Stratified CP
GT = Gap Time, M = Marginal

Table 8.12 presents, again for the exposure variable **tx** only, the estimated regression coefficients and robust standard errors for both the interaction and the no-interaction SC models for the three approaches (other than the CP approach) that use a SC model.

Notice that for each of the three SC modeling approaches, the estimated βs and corresponding standard errors are different over the four strata as well as for the no-interaction model. For example, using the **Stratified CP** approach, the estimated coefficients are −0.518, − **0.459,** − **0.117,** − **0.041,** and − **0.334** for strata 1 through 4 and the no-interaction model, respectively.

Version 1: Interaction SC Model

$h_g(t,\mathbf{X}) = h_{0g}(t)\mathbf{exp}[\beta_g\mathbf{tx} + \gamma_{1g}\,\mathbf{num} + \gamma_{2g}\,\mathbf{size}]$

$g = 1, 2, 3, 4$

Note: subscript g allows for different regression coefficients for each stratum

Such differing results over the different strata should be expected because they result from fitting an interaction SC model, which by definition allows for different regression coefficients over the strata.

Stratified CP for Subject 10

id	int	event	start	stop	tx	num	size
10	1	1	0	12	0	1	1

Gap Time for Subject 10

id	int	event	start	stop	tx	num	size
10	1	1	0	12	0	1	1

Marginal Approach for Subject 10

id	int	event	stime	tx	num	size
10	1	1	12	0	1	1

Note: int = stratum #

Notice also that for stratum 1, the estimated β and its standard error are identical (−**0.518** and **0.308**, resp.) for the **Stratified CP, Gap Time,** and **Marginal** approaches. This is as expected because, as illustrated for subject #10 at the left, the survival time information for first stratum is the same for stratum 1 for the three SC approaches and does not start to differ until stratum 2.

Marginal approach

start time = 0 always
stop time = **stime**

Subject # 10: (start, stop) = (0, 12)

Although the data layout for the **marginal** approach does not require (start,stop) columns, the start time for the first stratum (and all other strata) is 0 and the stop time is given in the **stime** column. In other words, for stratum 1 on subject #10, the stop time is 0 and the start time is 12, which is the same as for the **Stratified CP** and **Gap Time** data for this subject.

Bladder Cancer Study

1. Which approach is best?
2. Conclusion about **tx**?

So, based on all the information we have provided above concerning the analysis of the bladder cancer study,

1. Which of the four recurrent event analysis approaches is best?
2. What do we conclude about the estimated effect of **tx** controlling for num and size?

Which of the 4 approaches is best?
It depends!

The answer to question 1 is probably best phrased as, "**It depends!**" Nevertheless, if the investigator does not want to distinguish between recurrent events on the same subject and wishes an overall conclusion about the effect of **tx,** then the **CP** approach seems quite appropriate, as for this study.

CP: Don't want to distinguish recurrent event order
Want overall effect

If event order important:

Choose from the 3 SC approaches.

If, however, the investigator wants to distinguish the effects of **tx** according to the order that the event occurs (i.e., by stratum #), then one of the three SC approaches should be preferred. So, which one?

Stratified CP: time of recurrent event from entry into the study

The **Stratified CP** approach is preferred if the study goal is to use time of occurrence of each recurrent event from entry into the study to assess a subject's risk for an event of a specific order (i.e., as defined by a stratum #) to occur.

Gap Time: Use time from previous event to next recurrent event

The **Gap Time** approach would be preferred if the time interval of interest is the time (reset from 0) from the previous event to the next recurrent event rather than time from study entry until each recurrent event.

Marginal: Consider strata as representing different event types

Finally, the **Marginal** approach is recommended if the investigator wants to consider the events occurring at different orders as different types of events, for example different disease conditions.

Stratified CP versus **Marginal** (subtle choice)

Recommend: Choose Stratified CP unless strata represent different event types

We (the authors) consider the choice between the **Stratified CP** and **Marginal** approaches to be quite subtle. We prefer **Stratified CP,** provided the different strata do not clearly represent different event types. If, however, the strata clearly indicate separate event processes, we would recommend the **Marginal** approach.

What do we conclude about **tx**?

Conclusions based on results from **CP** and **Stratified CP** approaches

Overall, based on the above discussion, we think that the **CP** approach is an acceptable method to use for analyzing the bladder cancer data. If we had to choose one of the three SC approaches as an alternative, we would choose the **Stratified CP** approach, particularly because the order of recurrent events that define the strata doesn't clearly distinguish separate disease processes.

Table 8.13. Comparison of Results Obtained from No-Interaction Models Across Two Methods for Bladder Cancer Data

	Counting process	Stratified CP
Parameter estimate	−0.407	−0.334
Robust standard error	0.2418	0.1971
Wald chi-square	2.8338	2.8777
p-value	0.0923	0.0898
Hazard ratio	0.667	0.716
95% confidence interval	(0.414, 1.069)	(0.486, 1.053)

Table 8.13 summarizes the results for the **CP** and **Stratified CP** approaches with regard to the effect of the treatment variable (**tx**), adjusted for the control variables **num** and **size**. We report results only for the no-interaction models, because the interaction SC model for the **Stratified CP** approach was found (using LR test) to be not significantly different from the no-interaction model.

The results are quite similar for the two different approaches. There appears to be a small effect of **tx** on survival from bladder cancer: $\widehat{\text{HR}}(\textbf{CP}) = 0.667 = 1/1.50$, $\widehat{\text{HR}}(\textbf{C1}) = 0.716 = 1/1.40$. This effect is borderline nonsignificant (2-sided tests): P(**CP**) = .09 = P(**SCP**). 95% confidence intervals around the hazard ratios are quite wide, indicating an imprecise estimate of effect.

Overall, therefore, these results indicate that there is no strong evidence that **tx** is effective (after controlling for **num** and **size**) based on recurrent event survival analyses of the bladder cancer data.

X. A Parametric Approach Using Shared Frailty

Compared 4 approaches in previous section

- Each used a Cox model
- Robust standard errors
  - Adjusts for correlation from same subject

In the previous section we compared results obtained from using four analytic approaches on the recurrent event data from the bladder cancer study. Each of these approaches used a Cox model. Robust standard errors were included to adjust for the correlation among outcomes from the same subject.

We now present a parametric approach

- Weibull PH model
- Gamma shared frailty component
- Bladder Cancer dataset
  - Data layout for the counting process approach

Can review Chapter 7
Weibull model (Section VI)
Frailty models (Section XII)

In this section we present a parametric approach for analyzing recurrent event data that includes a frailty component. Specifically, a Weibull PH model with a gamma distributed shared frailty component is shown using the Bladder Cancer dataset. The data layout is the same as described for the counting process approach. It is recommended that the reader first review Chapter 7, particularly the sections on Weibull models (Section VI) and frailty models (Section XII).

Hazard conditioned on frailty α_k

$h_i(t|\alpha,\mathbf{X}_i) = \alpha_i h(t|\mathbf{X}_i)$

where $\alpha \sim$ gamma($\mu = 1$, var $= \theta$) and where $h(t|\mathbf{X}_i) = \lambda_i pt^{p-1}$ (Weibull) with $\lambda_{fk} = \exp(\beta_0 + \beta_1 \mathbf{tx}_i + \beta_2 \mathbf{num}_i + \beta_3 \mathbf{size}_i)$

We define the model in terms of the hazard of any (recurrent) outcome on the ith subject conditional on his or her frailty α_i. The frailty is a multiplicative random effect on the hazard function $h(t|\mathbf{X}_i)$, assumed to follow a gamma distribution of mean 1 and variance theta θ. We assume $h(t|\mathbf{X}_i)$ follows a Weibull distribution (shown at left).

Including shared frailty

- Accounts for unobserved factors
  - Subject specific
  - Source of correlation
  - Observations clustered by subject

The frailty is included in the model to account for variability due to unobserved subject-specific factors that are otherwise unaccounted for by the other predictors in the model. These unobserved subject-specific factors can be a source of within-subject correlation. We use the term **shared frailty** to indicate that observations are clustered by subject and each cluster (i.e., subject) shares the same level of frailty.

Robust standard errors

- Adjusts standard errors
- Does not affect coefficient estimates

In the previous sections, we have used robust variance estimators to **adjust** the standard errors of the coefficient estimates to account for within-subject correlation. Shared frailty is not only an adjustment, but also is built into the model and can have an impact on the estimated coefficients as well as their standard errors.

Shared frailty

- Built into model
- Can affect coefficient estimates and their standard errors

The model output (obtained using Stata version 10) is shown on the left. The inclusion of frailty in a model (shared or unshared) leads to one additional parameter estimate in the output (theta, the variance of the frailty). A likelihood ratio test for theta = 0 yields a statistically significant p-value of 0.003 (bottom of output) suggesting that the frailty component contributes to the model and that there is within-subject correlation.

Weibull regression (PH form)
Gamma shared frailty
Log likelihood = −184.73658

_t	Coef.	Std. Err.	z	P > |z|
tx	−.458	.268	−1.71	0.011
num	.184	.072	2.55	0.327
size	−.031	.091	−0.34	0.730
_cons	−2.952	.417	−7.07	0.000
/ln_p	−.119	.090	−1.33	0.184
/ln_the	−.725	.516	−1.40	0.160
p	.888	.080		
1/p	1.13	.101		
theta	.484	.250		

Likelihood ratio test of theta = 0:
chibar(01) = 7.34
Prob > = chibar2 = 0.003

The estimate for the Weibull shape parameter p is 0.888 suggesting a slightly decreasing hazard over time because $\hat{p} < 1$. However, the Wald test for $\ln(p) = 0$ (or equivalently $p = 1$) yields a non-significant p-value of 0.184.

Comparing Hazard Ratios

Weibull with frailty model

$\widehat{HR}(\mathbf{tx}) = \exp(-0.458) = 0.633$

$$95\%\ \text{CI} = \exp[-0.458 \pm 1.96(0.268)] = (0.374,\ 1.070)$$

An estimated hazard ratio of 0.633 for the effect of treatment comparing two individuals with the same level of frailty and controlling for the other covariates is obtained by exponentiating the estimated coefficient (−0.458) for **tx**. The estimated hazard ratio and 95% confidence intervals are similar to the corresponding results obtained using a counting processes approach with a Cox model and robust standard errors (see left).

Counting processes approach with Cox model

$\widehat{HR}(\mathbf{tx}):\ \exp(-0.407) = 0.667$

95% CI for HR **tx** (robust): (0.414, 1.069)

Interpretations of HR from frailty model

- Compares 2 individuals with same α
- Compares individual with himself
  - What is effect if individual had used treatment rather than placebo?

Another interpretation for the estimated hazard ratio from the frailty model involves the comparison of an individual to himself. In other words, this hazard ratio describes the effect on an individual's hazard (i.e., conditional hazard) if that individual had used the treatment rather than the placebo.

XI. A Second Example

Age-Related Eye Disease Study (AREDS)

Outcome

Age-related macular degeneration (AMD)

Clinical trial
Evaluate effect of treatment with high doses of antioxidants and zinc on progression of AMD

n = 43 (subset of data analyzed here)

We now illustrate the analysis of recurrent event survival data using a new example. We consider a subset of data from the Age-Related Eye Disease Study (AREDS), a long-term multicenter, prospective study sponsored by the U.S. National Eye Institute of the clinical course of age-related macular degeneration (AMD) (see AREDS Research Group, 2003).

In addition to collecting natural history data, AREDS included a clinical trial to evaluate the effect of high doses of antioxidants and zinc on the progression of AMD. The data subset we consider consists of 43 patients who experienced ocular events while followed for their baseline condition, macular degeneration.

Exposure

tx = 1 if treatment, 0 if placebo

8 years of follow-up

The exposure variable of interest was treatment group (**tx**), which was coded as 1 for patients randomly allocated to an oral combination of antioxidants, zinc, and vitamin C versus 0 for patients allocated to a placebo. Patients were followed for 8 years.

Two possible events

First event: visual acuity score <50 (i.e., poor vision)

Each patient could possibly experience two events. The first event was defined as the sudden decrease in visual acuity score below 50 measured at scheduled appointment times. Visual acuity score was defined as the number of letters read on a standardized visual acuity chart at a distance of 4 m, where the higher the score, the better the vision.

Second event: clinically advanced severe stage of macular degeneration

The second event was considered a successive stage of the first event and defined as a clinically advanced and severe stage of macular degeneration. Thus, the subject had to experience the first event before he or she could experience the second event.

4 approaches for analyzing recurrent event survival data carried out on macular degeneration data

Each model contains **tx, age,** and **sex**.

We now describe the results of using the four approaches for analyzing recurrent event survival with these data. In each analysis, two covariates **age** and **sex** were controlled, so that each model contained the variables **tx, age,** and **sex**.

CP model

$$h(t,\mathbf{X}) = h_0(t)\mathbf{exp}[\beta\ \mathbf{tx} + \gamma_1\ \mathbf{age} + \gamma_2\ \mathbf{sex}]$$

No-interaction SC model

$$h_g(t,\mathbf{X}) = h_{0g}(t)\mathbf{exp}[\beta\ \mathbf{tx} + \gamma_1\ \mathbf{age} + \gamma_2\ \mathbf{sex}]$$

where g = 1, 2

Interaction SC model:

$$h_g(t,\mathbf{X}) = h_{0g}(t)\mathbf{exp}[\beta_g\ \mathbf{tx} + \gamma_{1g}\ \mathbf{age} + \gamma_{2g}\ \mathbf{sex}]$$

where g = 1, 2

The counting process **(CP)** model is shown here at the left together with both the no-interaction and interaction SC models used for the three stratified Cox (SC) approaches.

Table 8.14 Comparison of Parameter Estimates and Robust Standard Errors for Treatment Variable (tx) Controlling for Age and Sex (Macular Degeneration Data)

	"Interaction" Cox stratified model		"No-interaction" SC model
Model	Stratum 1	Stratum 2	
	$\hat{\beta}_1$ (SE)	$\hat{\beta}_2$ (SE)	$\hat{\beta}_3$ (SE)
Counting process	n/a	n/a	−0.174 (0.104)
SCP	−0.055 (0.286)	−0.955 (0.443)	−0.306 (0.253)
GT	−0.055 (0.286)	−1.185 (0.555)	−0.339 (0.245)
Marginal	−0.055 (0.286)	−0.861 (0.465)	−0.299 (0.290)

In Table 8.14, we compare the coefficient estimates and their robust standard errors for the treatment variable (**tx**) from all four approaches. This table shows results for both the "interaction" and "nointeraction" stratified Cox models for the three approaches other than the counting process approach.

Notice that the estimated coefficients for β_1 and their corresponding standard errors are identical for the three SC approaches. This will always be the case for the first stratum regardless of the data set being considered.

The estimated coefficients for β_2 are, as expected, somewhat different for the three SC approaches. We return to these results shortly.

Interaction SC models are preferred (based on LR test results) to use of no-interaction SC model

LR tests for comparing the "no-interaction" with the "interaction" SC models were significant ($P < .0001$) for all three SC approaches (details not shown), indicating that an interaction model was more appropriate than a no-interaction model for each approach.

Table 8.15. Comparison of Results for the Treatment Variable (tx) Obtained for **Stratified CP** and Marginal Approaches (Macular Degeneration Data)

		Stratified CP	Marginal
Estimate	$\hat{\beta}_1$	−0.0555	−0.0555
	$\hat{\beta}_2$	−0.9551	−0.8615
	$\hat{\beta}$	−0.306	−0.2989
Robust std. error	$SE(\hat{\beta}_1)$	0.2857	0.2857
	$SE(\hat{\beta}_2)$	0.4434	0.4653
	$SE(\hat{\beta})$	0.2534	0.2902
Wald chi-square	$H_0{:}\beta_1 = 0$	0.0378	0.0378
	$H_0{:}\beta_2 = 0$	**4.6395**	**3.4281**
	$H_0{:}\beta = 0$	1.4569	1.0609
P-value	$H_0{:}\beta_1 = 0$	0.8458	0.8478
	$H_0{:}\beta_2 = 0$	**0.0312**	**0.0641**
	$H_0{:}\beta = 0$	0.2274	0.3030
Hazard ratio	$\exp(\hat{\beta}_1)$	0.946	0.946
	$\exp(\hat{\beta}_2)$	**0.385**	**0.423**
	$\exp(\hat{\beta})$	0.736	0.742
95% Conf. interval	$\exp(\hat{\beta}_1)$	(0.540, 1.656)	(0.540, 1.656)
	$\exp(\hat{\beta}_2)$	**(0.161, 0.918)**	**(0.170, 1.052)**
	$\exp(\hat{\beta})$	(0.448, 1.210)	(0.420, 1.310)

In Table 8.15, we summarize the statistical inference results for the effect of the treatment variable (**tx**) for the **Stratified CP** and **Marginal** approaches only.

We have not included the **CP** results here because the two events being considered are of very different types, particularly regarding severity of illness, whereas the CP approach treats both events as identical replications. We have not considered the **Gap Time** approach because the investigators were more likely interested in survival time from baseline entry into the study than the survival time "gap" from the first to second event.

Because we previously pointed out that the interaction SC model was found to be significant when compared to the corresponding no-interaction SC model, we focus here on the treatment (**tx**) effect for each stratum (i.e., event) separately.

First event:

	SCP	Marginal
$\widehat{HR}$	0.946	0.946
p-value	0.85	0.85

Based on the Wald statistics and corresponding P-values for testing the effect of the treatment on survival to the *first event* (i.e., H_0: $\beta_1 = 0$), both the **Stratified CP** and **Marginal** approaches give the identical result that the estimated treatment effect ($\widehat{HR} = 0.946 = 1/1.06$) is neither meaningful nor significant ($P = 0.85$).

Second event:

	SCP	Marginal
$\widehat{HR}$	0.385	0.423
p-value	0.03	0.06
95% CI	(0.16, 0.92)	(0.17, 1.05)

For the *second event*, indicating a clinically severe stage of macular degeneration, the Wald statistic P-value for the **Stratified CP** approach is 0.03, which is significant at the .05 level, whereas the corresponding P-value for the **Marginal** approach is 0.06, borderline nonsignificant at the .05 level.

The estimated HR for the effect of the treatment is ($\widehat{HR} = 0.385 = 1/2.60$) using the **Stratified CP** approach and its 95% confidence interval is quite wide but does not contain the null value of 1. For the **Marginal** approach, the estimated HR is $\widehat{HR} = 0.423 = 1/2.36$, also with a wide confidence interval, but includes 1.

Conclusions regarding 1st event:

- No treatment effect
- Same for **Stratified CP** and **Marginal** approaches

Thus, based on the above results, there appears to be no effect of treating patients with high doses of antioxidants and zinc on reducing visual acuity score below 50 (i.e., the first event) based on either **Stratified CP** or **Marginal** approaches to the analysis.

Conclusions regarding 2nd event:

- Clinically moderate and statistically significant treatment effect
- Similar for **Stratified CP** and **Marginal** approaches, but more support from **Stratified CP** approach

However, there is evidence of a clinically moderate and statistically significant effect of the treatment on protection (i.e., not failing) from the second more severe event of macular degeneration. This conclusion is more supported from the **Stratified CP** analysis than from the **Marginal** analysis.

Comparison of **Stratified CP** with **Marginal** Approach

What if results had been different?

Despite similar conclusions from both approaches, it still remains to compare the two approaches for these data. In fact, if the results from each approach had been very different, it would be important to make a choice between these approaches.

Recommend **Stratified CP** if

Can assume 2nd event cannot occur without 1st event previously occurring
⇓
Should consider survival time to 2nd event **conditional on** experiencing 1st event

Nevertheless, we authors find it difficult to make such a decision, even for this example. The **Stratified CP** approach would seem appropriate if the investigators assumed that the second event cannot occur without the first event previously occurring. If so, it would be important to consider survival time to the second event only for (i.e., **conditional on**) those subjects who experience a first event.

Recommend **Marginal** if

Can assume each subject at risk for 2nd event whether or not 1st event previously occurred
⇓
2nd event considered a separate event, that is, **unconditional** of the 1st event
⇓
Should consider survival times to 2nd event for **all** subjects

On the other hand, the **Marginal** approach would seem appropriate if each subject is considered to be at risk for the second event whether or not the subject experiences the first event. The second event is therefore considered separate from (i.e., **unconditional of**) the first event, so that survival times to the second event need to be included for all subjects, as in the **Marginal** approach.

Macular degeneration data: recommend **Marginal** approach

In general: carefully consider interpretation of each approach

For the macular degeneration data example, we find the **Marginal** approach persuasive. However, **in general**, the choice among all four approaches is not often clear-cut and requires careful consideration of the different interpretations that can be drawn from each approach.

XII. Survival Curves with Recurrent Events

Goal: Plot and Interpret Survival Curves

Types of survival curves:

KM (empirical): Chapter 2
Adjusted (Cox PH): Chapters 3 and 4

An important goal of most survival analyses, whether or not a regression model (e.g., Cox PH) is involved, is to plot and interpret/compare survival curves for different groups. We have previously described the Kaplan–Meier (KM) approach for plotting empirical survival curves (Chapter 2) and we have also described how to obtain adjusted survival curves for Cox PH models (Chapters 3 and 4).

Previously: 1 (nonrecurrent) event
Now:
Survival plots with recurrent events?

This previous discussion only considered survival data for the occurrence of one (nonrecurrent) event. So, how does one obtain survival plots when there are recurrent events?

Focus on one ordered event at a time

$S_1(t)$: 1st event
$S_2(t)$: 2nd event
...
$S_k(t)$: kth event

The answer is that survival plots with recurrent events only make sense when the focus is on one ordered event at a time. That is, we can plot a survival curve for **survival to a first event, survival to a second event,** and so on.

Survival to a 1st event

$S_1(t) = \Pr(T_1 > t)$

where
T_1 = survival time up to occurrence of 1st event
(ignores later recurrent events)

For **survival to a first event,** the survival curve describes the probability that a subject's time to occurrence of a first event will exceed a specified time. Such a plot essentially ignores any recurrent events that a subject may have after a first event.

Survival to a 2nd event

$S_2(t) = \Pr(T_2 > t)$

where
T_2 = survival time up to occurrence of 2nd event

For **survival to a second event**, the survival curve describes the probability that a subject's time to occurrence of a second event will exceed a specified time.

Two versions

There are two possible versions for such a plot.

Stratified:
T_{2c} = time from 1st event to 2nd event, restricting data to 1st event subjects
Marginal:
T_{2m} = time from study entry to 2nd event, ignoring 1st event

Stratified: use survival time from time of first event until occurrence of second event, thus restricting the dataset to only those subjects who experienced a first event.
Marginal: use survival time from study entry to occurrence of second event, ignoring whether a first event occurred.

Survival to a kth event (k ≥ 2)

$S_k(t) = \Pr(T_k > t)$

where
T_k = survival time up to occurrence of kth event

Similarly, for **survival to the kth event**, the survival curve describes the probability that a subject's time to occurrence of the kth event will exceed a specified time.

Two versions

Stratified:
T_{kc} = time from the $k - 1$st to kth event, restricting data to subjects with $k - 1$ events
Marginal:
T_{km} = time from study entry to kth event, ignoring previous events

As with survival to the second event, there are two possible versions, **Stratified** or **Marginal,** for such a plot, as stated on the left.

EXAMPLE

			Days		
ID	**Status**	**Stratum**	**Start**	**Stop**	**tx**
M	1	1	0	100	1
M	1	2	100	105	1
H	1	1	0	30	0
H	1	2	30	50	0
P	1	1	0	20	0
P	1	2	20	60	0
P	1	3	60	85	0

We now illustrate such survival plots for recurrent event data by returning to the small dataset previously described for three subjects Molly (M), Holly (H), and Polly (P), shown again on the left.

Deriving $S_1(t)$: **Stratum 1**

$t_{(f)}$	n_f	m_f	q_f	$R(t_{(f)})$	$S_1(t_{(f)})$
0	3	0	0	{M, H, P}	1.00
20	3	1	0	{M, H, P}	0.67
30	2	1	0	{M, H}	0.33
100	1	1	0	{M}	0.00

The survival plot for survival to the first event $\mathbf{S_1(t)}$ is derived from the stratum 1 data layout for any of the three alternative SC analysis approaches. Recall that m_f and q_f denote the number of failures and censored observations at time $t_{(f)}$. The survival probabilities in the last column use the KM product limit formula.

Deriving $S_{2c}(t)$: **Stratum 2 (Stratified GT)**

$t_{(f)}$	n_f	m_f	q_f	$R(t_{(f)}\}$	$S_{2c}(t_{(f)})$
0	3	0	0	{M, H, P}	1.00
5	3	1	0	{M, H, P}	0.67
20	2	1	0	{M, P}	0.33
450	1	1	0	{M}	0.00

The **Stratified** survival plot for survival to the second event is derived from the stratum 2 data layout for the **Gap Time** approach. We denote this survival curve as $\mathbf{S_{2c}(t)}$. Notice that the survival probabilities here are identical to those in the previous table; however, the failure times $t_{(f)}$ in each table are different.

Deriving $S_{2m}(t)$: **Stratum 2 (Marginal)**

$t_{(f)}$	n_f	m_f	q_f	$R(t_{(f)}\}$	$S_{2m}(t_{(f)})$
0	3	0	0	{M, H, P}	1.00
20	3	1	0	{M, H, P}	0.67
30	2	1	0	{H, P}	0.33
100	1	1	0	{P}	0.00

The **Marginal** survival plot for survival to the second event is derived from the stratum 2 data layout for the **Marginal** approach. We denote this survival curve as $\mathbf{S_{2m}(t)}$. Again, the last column here is identical to those in the previous two tables, but, once again, the failure times $t_{(f)}$ in each table are different.

Survival Plots for Molly, Holly and Polly Recurrent Event Data (n = 3)

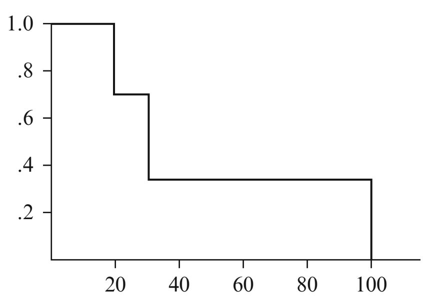


Figure 8.1. $\mathbf{S_1(t)}$: Survival to 1st Event

The survival plots that correspond to the above three data layouts are shown in Figures 8.1 to 8.3.

Figure 8.1 shows survival probabilities for the first event, ignoring later events. The risk set at time zero contains all three subjects. The plot drops from $S_1(t) = 1$ to $S_1(t) = 0.67$ at $t = 20$, drops again to $S_1(t) = 0.33$ at $t = 30$ and falls to $S_1(t) = 0$ at $t = 100$ when the latest first event occurs.

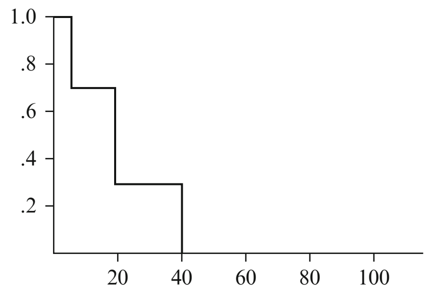


Figure 8.2. $\mathbf{S}_{2c}(t)$: Survival to 2nd Event (Stratified GT)

Figure 8.2 shows **Stratified GT** survival probabilities for the second event using survival time **from the first event to the second event**. Because all three subjects had a first event, the risk set at time zero once again contains all three subjects. Also, the survival probabilities of 1, 0.67, 0.33, and 0 are the same as in Figure 8.1. Nevertheless, this plot differs from the previous plot because the survival probabilities are plotted at different survival times ($t = 5, 20, 40$ in Figure 8.2 instead of $t = 20, 30, 100$ in Figure 8.1)

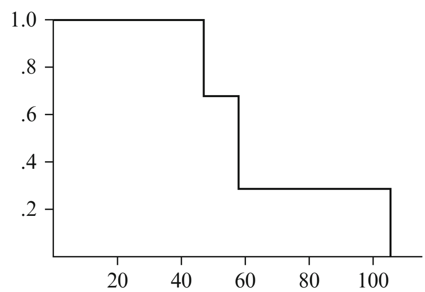


Figure 8.3. $\mathbf{S}_{2m}(t)$: Survival to 2nd Event (Marginal)

Figure 8.3 shows **Marginal** survival probabilities for the second event using survival time **from study entry to the second event, ignoring the first event**. The survival probabilities of 1, 0.67, 0.33, and 0 are once again the same as in Figures 8.1 and 8.2. Nevertheless, this plot differs from the previous two plots because the survival probabilities are plotted at different survival times ($t = 50, 60, 105$ in Figure 8.3).

XIII. Summary

We have described four approaches for analyzing recurrent event survival data.

4 approaches for recurrent event data
Counting process (CP), Stratified CP, Gap Time, Marginal

These approaches differ in how the risk set is determined and in data layout. All four approaches involve using a standard computer program that fits a Cox PH model, with the latter three approaches requiring a stratified Cox model, stratified by the different events that occur.

The 4 approaches

- Differ in how risk set is determined
- Differ in data layout
- All involve standard Cox model program
- Latter three approaches use a SC model

Identical recurrent events
⇓
CP approach

The approach to analysis typically used when recurrent events are treated as identical is called the **CP Approach**.

Recurrent events: different disease categories or event order important
⇓
Stratified Cox (SC) approaches

When recurrent events involve different disease categories and/or the order of events is considered important, the analysis requires choosing among the three alternative SC approaches.

CP approach: **Start** and **Stop** times

Standard layout: only **Stop** (survival) times (no recurrent events)

The data layout for the counting process approach requires each subject to have a line of data for each recurrent event and lists the **start time** and **stop time** of the interval of follow-up. This contrasts with the standard layout for data with no recurrent events, which lists only the stop (**survival**) time on a single line of data for each subject.

Stratified CP: same **Start** and **Stop** Times as **CP**, but uses **SC** model

The **Stratified CP** approach uses the exact same (start, stop) data layout format used for the **CP** approach, except that for **Stratified CP,** the model used is a SC PH model rather than an unstratified PH model.

Gap Time: **Start** and **Stop** Times
Start = 0 always
Stop = time since previous event
SC model

The **Gap Time** approach also uses a (start, stop) data layout, but the start value is always 0 and the stop value is the time interval length since the previous event. The model here is also a SC model.

Marginal approach:

Standard layout (nonrecurrent event), that is, without (**Start, Stop**) columns
Each failure is a separate process

The **Marginal** approach uses the standard (nonrecurrent event) data layout instead of the (start, stop) layout. The basic idea behind the **Marginal** approach is that it allows each failure to be considered as a separate process.

Recommend using **robust estimation** to adjust for correlation of observations on the same subject.

For each of the SC alternative approaches, as for the **CP** approach, it is recommended to use **robust estimation** to adjust the variances of the estimated regression coefficients for the correlation of observations on the same subject.

Application 1: Bladder Cancer study
n = 86
64 months of follow-up

We considered two applications of the different approaches described above. First, we compared results from using all four methods to analyze data from a study of bladder cancer involving 86 patients, each followed for a variable time up to 64 months.

Repeated event: recurrence of bladder cancer tumor; up to 4 events

The repeated event analyzed was the recurrence of a bladder cancer tumor after transurethral surgical excision. Each recurrence of new tumors was treated by removal at each examination. About 25% of the 86 subjects experienced four events.

tx = 1 if thiotepa, 0 if placebo
num = initial # of tumors
size = initial size of tumors

The exposure variable of interest was drug treatment status (**tx,** 0 = placebo, 1 = treatment with thiotepa), There were two covariates: initial number of tumors (**num**) and initial size of tumors (**size**).

CP results: no strong evidence for **tx** ($\widehat{HR}$ = 0.67, P = .09, 95% CI: 0.414, 1.069)

Results for the **CP** approach, which was considered appropriate for these data, indicated that there was no strong evidence that **tx** is effective after controlling for **num** and **size**.

Alternative parametric approach

- Weibull PH model
- Gamma shared frailty component
- Bladder cancer dataset
- Similar HR and confidence interval as for counting process approach

An alternative approach for analyzing recurrent event data was also described using a parametric model containing a frailty component (see Chapter 7). Specifically, a Weibull PH model with a gamma distributed frailty was fit using the bladder cancer dataset. The resulting estimated HR and confidence interval were quite similar to the counting process results.

Application 2: Clinical trial

n = 43
8 years of follow-up
High doses of antioxidants and zinc
Age-related macular degeneration

The second application considered a subset of data (n = 43) from a clinical trial to evaluate the effect of high doses of antioxidants and zinc on the progression of age-related macular degeneration (AMD). Patients were followed for 8 years.

Exposure: **tx** = 1 if treatment, 0 if placebo
Covariates: **age, sex**

The exposure variable of interest was treatment group (**tx**). Covariates considered were **age** and **sex**.

Two possible events:

1st event: visual acuity score <50 (i.e., poor vision)
2nd event: clinically advanced severe stage of macular degeneration

Each patient could possibly experience two events. The first event was defined as the sudden decrease in visual acuity score below 50. The second event was considered a successive stage of the first event and defined as a clinically advanced and severe stage of macular degeneration.

Focus on **Stratified CP** vs. **Marginal** (events were of different types)

Because the two events were of very different types and because survival from baseline was of primary interest, we focused on the results for the **Stratified CP** and **Marginal** approaches only.

Interaction SC model ✓
No-interaction SC model ×

An interaction SC model was more appropriate than a no-interaction model for each approach, thus requiring separate results for the two events under study.

Conclusions regarding 1st event

- No treatment effect
- Same for **Stratified CP** and **Marginal** approaches

The results for the first event indicated no effect of the treatment on reducing visual acuity score below 50 (i.e., the first event) from either **Stratified CP** or **Marginal** approaches to the analysis.

Conclusions regarding 2nd event

- Clinically moderate and statistically significant treatment effect

However, there was evidence of a clinically moderate and statistically significant effect of the treatment on the second more severe event of macular degeneration.

Macular degeneration data: prefer **Marginal** approach (but not clear-cut)

The choice between the Stratified CP and marginal approaches for these data was not clear-cut, although the **Marginal** approach was perhaps more appropriate because the two events were of very different types.

In general: carefully consider interpretation of each approach

In general, however, the choice among all four approaches requires careful consideration of the different interpretations that can be drawn from each approach.

Survival plots: one ordered event at a time Two versions for survival to kth event:
Stratified: only subjects with $k - 1$ events
Marginal: ignores previous events

Survival plots with recurrent events are derived one ordered event at a time. For plotting survival to a kth event where $k \geq 2$, one can use either a **Stratified** or **Marginal** plot, which typically differ.

Detailed Outline

I. Overview (page 366)

A. Focus: outcome events that may occur more than once over the follow-up time for a given subject, that is, "recurrent events."

B. **Counting Process (CP)** approach uses the Cox PH model.

C. Alternative approaches that use a Stratified Cox (SC) PH model and a frailty model.

II. Examples of Recurrent Event Data (pages 366–368)

A. 1. Multiple relapses from remission: leukemia patients.
   2. Repeated heart attacks: coronary patients.
   3. Recurrence of tumors: bladder cancer patients.
   4. Deteriorating episodes of visual acuity: macular degeneration patients.

B. Objective of each example: to assess relationship of predictors to rate of occurrence, allowing for multiple events per subject.

C. Different analysis required depending on whether:
   1. Recurrent events are treated as identical (**counting process approach**), or
   2. Recurrent events involve different disease categories and/or the order of events is important (**stratified Cox** approaches).

III. Counting Process Example (pages 368–369)

A. Data on two hypothetical subjects from a randomized trial that compares two treatments for bladder cancer tumors.

B. Data set-up for **Counting Process (CP)** approach:
   1. Each subject contributes a line of data for each time interval corresponding to each recurrent event and any additional event-free follow-up interval.
   2. Each line of data for a given subject lists the **start time** and **stop time** for each interval of follow-up.

IV. General Data Layout: Counting Process Approach (pages 370–371)

A. r_i time intervals for subject i.
d_{ij} event status (0 or 1) for subject i in interval j.
t_{ij0} start time for subject i in interval j.
t_{ij1} stop time for subject i in interval j.
X_{ijk} value of kth predictor for subject i in interval j.
$i = 1, 2, \ldots, N$; $j = 1, 2, \ldots, r_i$; $k = 1, 2, \ldots, p$.

B. Layout for subject i:

i	j	d_{ij}	t_{ij0}	t_{ij1}	X_{ij1}	X_{ijp}
i	1	d_{i1}	t_{i10}	t_{i11}	X_{111}	X_{i1p}
i	2	d_{i2}	t_{i20}	t_{i21}	X_{121}	X_{i2p}
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
i	r_i	d_{ir_i}	t_{ir_i0}	t_{ir_i1}	X_{ir_i1}	X_{ir_ip}

C. Bladder Cancer Study example:
1. Data layout provided for the first 26 subjects (86 subjects total) from a 64-month study of recurrent bladder cancer tumors.
2. The exposure variable: drug treatment status (**tx,** 0 = placebo, 1 = treatment with thiotepa).
3. Covariates: initial number of tumors (**num**) and initial size of tumors (**size**).
4. Up to 4 events per subject.

V. The Counting Process Model and Method (pages 372–376)

A. The model typically used to carry out the **Counting Process (CP)** approach is the standard Cox PH model: $h(t, \mathbf{X}) = h_0(t) \exp[\Sigma\ \beta_i X_i]$.

B. For recurrent event survival data, the (partial) likelihood function is formed differently than for nonrecurrent event survival data:
1. A subject who continues to be followed after having failed at $t_{(f)}$ does not drop out of the risk set after $t_{(f)}$ and remains in the risk set until his or her last interval of follow-up, after which the subject is removed from the risk set.
2. Different lines of data contributed by the same subject are treated in the analysis as if they were independent contributions from different subjects.

C. For the bladder cancer data, the Cox PH Model for **CP** approach is given by

$$h(t, \mathbf{X}) = h_0(t)\mathbf{exp}[\beta\ \mathbf{tx} + \gamma_1\ \mathbf{num} + \gamma_2\ \mathbf{size}].$$

D. The overall partial likelihood **L** from using the **CP** approach will be automatically determined by the computer program used once the data layout is in the correct **CP** form and the program code used involves the (start, stop) formulation.

VI. Robust Estimation (pages 376–378)

A. In the **CP** approach, the different intervals contributed by a given subject represent correlated observations on the same subject that must be accounted for in the analysis.

B. A widely used technique for adjusting for the correlation among outcomes on the same subject is called **robust estimation**.

C. The goal of **robust estimation** for the CP approach is to obtain variance estimators that adjust for correlation within subjects when previously no such correlation was assumed.

D. The **robust estimator** of the variance of an estimated regression coefficient allows tests of hypotheses and confidence interval estimation about model parameters to account for correlation within subjects.

E. The general form of the **robust estimator** can be most conveniently written in matrix notation; this formula is incorporated into the computer program and is automatically calculated by the program with appropriate coding.

VII. Results for CP Example (pages 378–379)

A. Edited output is provided from fitting the no-interaction Cox PH model involving the three predictors **tx, num**, and **size**.

B. A likelihood ratio chunk test for interaction terms **tx** × **num** and **tx** × **size** was nonsignificant.

C. The PH assumption was assumed satisfied for all three variables.

D. The robust estimator of 0.2418 for the standard deviation of **tx** was similar though somewhat different from the corresponding nonrobust estimator of 0.2001.

E. There was not strong evidence that **tx** is effective after controlling for **num** and **size** ($\widehat{HR} =$ 0.67, two-sided P = .09, 95% CI: 0.414, 1.069).

F. However, for a one-sided alternative, the p-values using both robust and nonrobust standard errors were significant at the .05 level.

G. The 95% confidence interval using the robust variance estimator is quite wide.

VIII. Other Approaches Stratified Cox
(pages 379–385)

A. The "strata" variable for each of the three SC approaches treats the time interval number for each event occurring on a given subject as a stratified variable.

B. Three alternative approaches involving SC models need to be considered if the investigator wants to distinguish the order in which recurrent events occur.

C. These approaches all differ from what is called **competing risk** survival analysis in that the latter allows each subject to experience only one of several different types of events over follow-up.

D. **Stratified CP** approach:
   1. Same **Start** and **Stop** Times as **CP**.
   2. **SC** model.

E. **Gap Time** approach:
   1. **Start** and **Stop** Times, but **Start** = 0 always and **Stop** = time since previous event.
   2. **SC** model.

F. **Marginal** approach:
   1. Uses standard layout (nonrecurrent event); no (**Start, Stop**) columns.
   2. Treats each failure is a separate process.
   3. Each subject at risk for all failures that might occur, so that # actual failures < # possible failures.
   4. **SC** model.

G. Must decide between two types of **SC** models:
   1. No-interaction **SC** versus interaction **SC**.
   2. Bladder cancer example:
      No-interaction model: $h_g(t, \mathbf{X}) = h_{0g}(t)\mathbf{exp}[\beta\ \mathbf{tx} + \gamma_1\ \mathbf{num} + \gamma_2\ \mathbf{size}]$ where g = 1, 2, 3, 4.
      Interaction model: $h_g(t,\mathbf{X}) = h_{0g}(t)\mathbf{exp}[\beta_g\mathbf{tx} + \gamma_{1g}\mathbf{num} + \gamma_{2g}\mathbf{size}]$. where g = 1, 2, 3, 4.

H. Recommend using **robust estimation** to adjust for correlation of observations on the same subject.

IX. Bladder Cancer Study Example (Continued) (pages 385–389)

A. Results from using all four methods **– CP, Stratified CP, Gap Time**, and **Marginal –** on the bladder cancer data were compared.

B. The hazard ratio for the effect of **tx** based on a no-interaction model differed somewhat for each of the four approaches, with the marginal model being most different:

M: 0.560 **CP:** 0.666 **SCP:** 0.716 **GT:** 0.763

C. The nonrobust and robust standard errors and P-values differed to some extent for each of the different approaches.

D. Using an interaction SC model, the estimated βs and corresponding standard errors are different over the four strata (i.e., four events) for each model separately.

E. The estimated β's and corresponding standard errors for the three alternative SC models are identical, as expected (always for first events).

F. Which of the four recurrent event analysis approaches is best?

1. Recommend **CP** approach if do not want to distinguish between recurrent events on the same subject and desire overall conclusion about the effect of **tx**.
2. Recommend one of the three SC approaches if want to distinguish the effect of **tx** according to the order in which the event occurs.
3. The choice between the **Stratified CP** and **Marginal** is difficult, but prefer **Stratified CP** because the strata do not clearly represent different event types.

G. Overall, regardless of the approach used, there was no strong evidence that **tx** is effective after controlling for **num** and **size**.

X. A Parametric Approach Using Shared Frailty (pages 389–391)

A. Alternative approach using a parametric model containing a frailty component (see Chapter 7).

B. Weibull PH model with a gamma distributed frailty was fit using the bladder cancer dataset.

C. Estimated HR and confidence interval were quite similar to the counting process results.

D. Estimated frailty component was significant ($P = 0.003$).

XI. A Second Example (pages 391–395)

A. Clinical trial (n = 43, 8-year study) on effect of using high doses of antioxidants and zinc (i.e., **tx** = 1 if yes, 0 if no) to prevent age-related macular degeneration.

B. Covariates: **age** and **sex**.

C. Two possible events:

1. First event: visual acuity score <50 (i.e., poor vision).
2. Second event: clinically advanced stage of macular degeneration.

D. Focus on **Stratified CP** vs. **Marginal** because events are of different types.

E. Interaction SC model significant when compared to no-interaction SC model.

F. Conclusions regarding 1st event:

1. No treatment effect (HR = 0.946, P = 0.85).
2. Same for **Stratified CP** and **Marginal** approaches.

G. Conclusions regarding 2nd event.

1. **Stratified CP:** $\widehat{HR} = 0.385 = 1/2.60$, two-sided P = 0.03.
2. **Marginal:** $\widehat{HR} = 0.423 = 1/2.36$, two-sided P = 0.06).
3. Overall, clinically moderate and statistically significant treatment effect.

H. **Marginal** approach preferred because 1st and 2nd events are different types.

XII. Survival Curves with Recurrent Events (pages 395–398)

A. Survival plots with recurrent events only make sense when the focus is on one ordered event at a time.

B. For survival from a 1st event, the survival curve is given by $S_1(t) = Pr\ (T_1 > t)$ where T_1 = survival time up to occurrence of the 1st event (ignores later recurrent events).

C. For survival from the kth event, the survival curve is given by $S_k(t) = Pr\ (T_k > t)$ where T_k = survival time up to occurrence of the kth event).

D. Two versions for $S_k(t)$:
   i. **$S_{kc}(t)$ Stratified:** T_{kc} = time from the (k-1)st to kth event, restricting data to subjects with k-1 events.
   ii. **$S_{km}(t)$ Marginal:** T_{km} = time from study entry to kth event, ignoring previous events.

E. Illustration of survival plots for recurrent event data using a small dataset involving three subjects Molly (M), Holly (H), and Polly (P).

XIII. Summary (pages 398–401)

A. Four approaches for analyzing recurrent event survival data: the **counting process (CP), Stratified CP, Gap Time**, and **Marginal** approaches.

B. Data layouts differ for each approach.

C. **CP** approach uses Cox PH model; other approaches use Cox SC model.

D. Choice of approach depends in general on carefully considering the interpretation of each approach.

E. Should use **robust estimation** to adjust for correlation of observations on the same subject.

Practice Exercises

Answer questions 1 to 15 as true or false (circle T or F).

T F 1. A recurrent event is an event (i.e., failure) that can occur more than once over the follow-up on a given subject.

T F 2. The **Counting Process (CP)** approach is appropriate if a given subject can experience more than one different type of event over follow-up.

T F 3. In the data layout for the **CP** approach, a subject who has additional follow-up time after having failed at time $t_{(f)}$ does not drop out of the risk set after $t_{(f)}$.

T F 4. The **CP** approach requires the use of a stratified Cox (SC) PH model.

T F 5. Using the **CP** approach, if exactly two subjects fail at month $t_{(f)} = 10$, but both these subjects have later recurrent events, then the number in the risk set at the next ordered failure time does not decrease because of these two failures.

T F 6. The goal of **robust estimation** for the **CP** approach is to adjust estimated regression coefficients to account for the correlation of observations within subjects when previously no such correlation was assumed.

T F 7. **Robust estimation** is recommended for the **CP** approach but not for the alternative SC approaches for analyzing recurrent event survival data.

T F 8. The p-value obtained from using a robust standard error will always be larger than the corresponding p-value from using a nonrobust standard error.

T F 9. The **Marginal** approach uses the exact same (start, stop) data layout format used for the **CP** approach, except that for the **Marginal** approach, the model used is a stratified Cox PH model variable rather than a standard (unstratified) PH model.

T F 10. Suppose the maximum number of failures occurring for a given subject is **five** in a dataset to be analyzed using the **Marginal** approach. Then a subject who failed only twice will contribute five lines of data corresponding to his or her two failures and the three additional failures that could have possibly occurred for this subject.

T F 11. Suppose the maximum number of failures occurring for a given subject is **five** in a dataset to be analyzed using the **Stratified CP** approach. Then an interaction SC model used to carry out this analysis will have the following general model form: $h_g(t, \mathbf{X}) = h_{0g}(t) \exp[\beta_{1g}X_1 + \beta_{2g}X_2 + \cdots + \beta_{pg}X_p]$, $g = 1,2,3,4,5$.

T F 12. Suppose a no-interaction SC model using the **Stratified CP** approach is found (using a likelihood ratio test) not statistically different from a corresponding interaction SC model. Then if the no-interaction model is used, it will not be possible to separate out the effects of predictors within each stratum representing the recurring events on a given subject.

T F 13. In choosing between the **Stratified CP** and the **Marginal** approaches, the **Marginal** approach would be preferred provided the different strata clearly represent different event types.

T F 14. When using an interaction SC model to analyze recurrent event data, the estimated regression coefficients and corresponding standard errors for the first stratum always will be identical for the **Stratified CP, Gap Time,** and **Marginal** approaches.

T F 15. The choice among the **CP, Stratified CP, Gap Time**, and **Marginal** approaches depends upon whether a no-interaction SC or an interaction SC model is more appropriate for one's data.

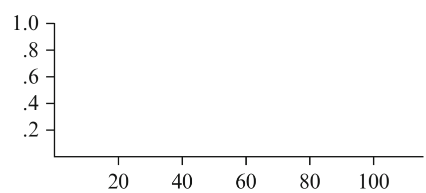
16. Suppose that Allie (A), Sally (S), and Callie (C) are the only three subjects in the dataset shown below. All three subjects have two recurrent events that occur at different times.

ID	Status	Stratum	Start	Stop	tx
A	1	1	0	70	1
A	1	2	70	90	1
S	1	1	0	20	0
S	1	2	20	30	0
C	1	1	0	10	1
C	1	2	10	40	1

Fill in the following data layout describing survival (in weeks) **to the first event (stratum 1)**. Recall that m_f and q_f denote the number of failures and censored observations at time $t_{(f)}$. The survival probabilities in the last column use the KM product limit formula.

$t_{(f)}$	n_f	m_f	q_f	$R(t_{(f)})$	$S_1(t_{(f)})$
0	3	0	0	{A, S, C}	1.00
10	-	-	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-

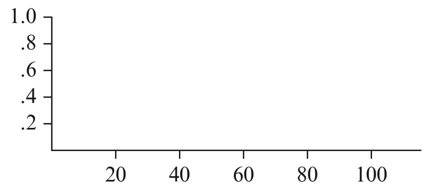
17. Plot the survival curve that corresponds to the data layout obtained for Question 16.



18. Fill in the following data layout describing survival (in weeks) **from the first to second event** using the **Gap Time** approach:

$t_{(f)}$	n_f	m_f	q_f	$R(t_{(f)})$	$S_2(t_{(f)})$
0	3	0	0	{A, S, C}	1.00
10	-	-	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-

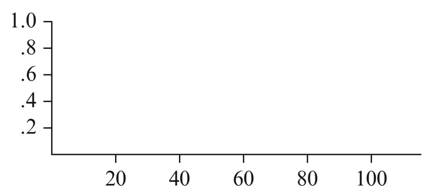
19. Plot the survival curve that corresponds to the data layout obtained for Question 18.



20. Fill in the following data layout describing survival (in weeks) **to the second event** using the **Marginal** approach:

$t_{(f)}$	n_f	m_f	q_f	$R(t_{(f)})$	$S_2(t_{(f)})$
0	3	0	0	{A, S, C}	1.00
30	-	-	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-

21. Plot the survival curve that corresponds to the data layout obtained for Question 20.



22. To what extent do the three plots obtained in Questions 17, 19, and 21 differ? Explain briefly.

Test

1. Suppose that Bonnie (B) and Lonnie (L) are th only two subjects in the dataset shown below, wher both subjects have two recurrent events that occur ar different times.

ID	Status	Stratum	Start	Stop
B	1	1	0	12
B	1	2	12	16
L	1	1	0	20
L	1	2	20	23

a. Fill in the empty cells in the following data layout describing survival time (say, in weeks) to the first event (stratum 1):

$t_{(f)}$	n_f	m_f	q_f	$R(t_{(f)})$
0	2	0	0	{B, L}
12				
20				

b. Why will the layout given in part a be the same regardless of whether rhe analysis approach is the Counting Process (CP), Stratified CP, Gap Time, or Marginal approaches?

c. Fill in the empty cells in the following data layout describing survival time (say, in weeks) from the first to the second event (stratum 2) using the Stratified CP approach:

$t_{(f)}$	n_f	m_f	q_f	$R(t_{(f)})$
0	0	0	0	-
16				
23				

d. Fill in the empty cells in the following data layout describing survival time (say, in weeks) from the first to the second event (stratum 2) using the Gap Time approach:

$t_{(f)}$	n_f	m_f	q_f	$R(t_{(f)})$
0	2	0	0	{B, L}
3				
4				

e. Fill in the empty cells in the following data layout describiing survival time (say, in weeks) from the first to the second event (stratum 2) using the Marginal approach:

$t_{(f)}$	n_f	m_f	q_f	$R(t_{(f)})$
0	2	0	0	{B, L}
16				
23				

f. For the **Stratified CP** approach described in part c, determine which of the following choices is correct. Circle the number corresponding to the one and only one correct choice.
   i. Lonnie is in the risk set when Bonnie gets her second event.
   ii. Bonnie is in the risk set when Lonnie gets her second event.
   iii. Neither is in the risk set for the other's second event.

g. For the **Gap Time** approach described in part d, determine which of the following choices is correct. Circle the number corresponding to the one and only one correct choice.
   i. Lonnie is in the risk set when Bonnie gets her second event.
   ii. Bonnie is in the risk set when Lonnie gets her second event.
   ii. Neither is in the risk set for the other's second event.

h. For the **Marginal** approach described in part e, determine which of the following choices is correct. Circle the number corresponding to the one and only one correct choice.
   i. Lonnie is in the risk set when Bonnie gets her second event.
   ii. Bonnie is in the risk set when Lonnie gets her second event.
   iii. Neither is in the risk set for the other's second event.

2. The dataset shown below in the counting process layout comes from a clinical trial involving 36 heart attack patients between 40 and 50 years of age with implanted defibrillators who were randomized to one of two treatment groups (**tx**, = 1 if treatment A, = 0 if treatment B) to reduce their risk for future heart attacks over a 4-month period. The event of interest was experiencing a "high energy shock" from the defibrillator. The outcome is time (in days) until an event occurs. The covariate of interest was Smoking History (1 = ever smoked, 0 = never smoked). Questions about the analysis of this dataset follow.

Col 1 = id, Col 2 = event, Col 3 = start, Col 4 = stop, Col 5 = tx, Col 6 = smoking

01	1	0	39	0	0	12	1	0	39	0	1
01	1	39	66	0	0	12	1	39	80	0	1
01	1	66	97	0	0	12	0	80	107	0	1
02	1	0	34	0	1	13	1	0	36	0	1
02	1	34	65	0	1	13	1	36	64	0	1
02	1	65	100	0	1	13	1	64	95	0	1
03	1	0	36	0	0	14	1	0	46	0	1
03	1	36	67	0	0	14	1	46	77	0	1
03	1	67	96	0	0	14	0	77	111	0	1
04	1	0	40	0	0	15	1	0	61	0	1
04	1	40	80	0	0	15	1	61	79	0	1
04	0	80	111	0	0	15	0	79	111	0	1
05	1	0	45	0	0	16	1	0	57	0	1
05	1	45	68	0	0	16	0	57	79	0	1
05	.	68	.	0	0	16	.	79	.	0	1
06	1	0	33	0	1	17	1	0	37	0	1
06	1	33	66	0	1	17	1	37	76	0	1
06	1	66	96	0	1	17	0	76	113	0	1
07	1	0	34	0	1	18	1	0	58	0	1
07	1	34	67	0	1	18	1	58	67	0	1
07	1	67	93	0	1	18	0	67	109	0	1
08	1	0	39	0	1	19	1	0	58	1	1
08	1	39	72	0	1	19	1	58	63	1	1
08	1	72	102	0	1	19	1	63	106	1	1
09	1	0	39	0	1	20	1	0	45	1	0
09	1	39	79	0	1	20	1	45	72	1	0
09	0	79	109	0	1	20	1	72	106	1	0
10	1	0	36	0	0	21	1	0	48	1	0
10	1	36	65	0	0	21	1	48	81	1	0
10	1	65	96	0	0	21	1	81	112	1	0
11	1	0	39	0	0	22	1	0	38	1	1
11	1	39	78	0	0	22	1	38	64	1	1
11	1	78	108	0	0	22	1	64	97	1	1

(*Continued on next page*)

23	1	0	51	1	1	30	1	0	57	1	0
23	1	51	69	1	1	30	1	57	78	1	0
23	0	69	98	1	1	30	1	78	99	1	0
24	1	0	43	1	1	31	1	0	44	1	1
24	1	43	67	1	1	31	1	44	74	1	1
24	0	67	111	1	1	31	1	74	96	1	1
25	1	0	46	1	0	32	1	0	38	1	1
25	1	46	66	1	0	32	1	38	71	1	1
25	1	66	110	1	0	32	1	71	105	1	1
26	1	0	33	1	1	33	1	0	38	1	1
26	1	33	68	1	1	33	1	38	64	1	1
26	1	68	96	1	1	33	1	64	97	1	1
27	1	0	51	1	1	34	1	0	38	1	1
27	1	51	97	1	1	34	1	38	63	1	1
27	0	97	115	1	1	34	1	63	99	1	1
28	1	0	37	1	0	35	1	0	49	1	1
28	1	37	79	1	0	35	1	49	70	1	1
28	1	79	93	1	0	35	0	70	107	1	1
29	1	0	41	1	1	36	1	0	34	1	1
29	1	41	73	1	1	36	1	34	81	1	1
29	0	73	111	1	1	36	1	81	97	1	1

Table T.1 below provides the results for the treatment variable (**tx**) from no-interaction models over all four recurrent event analysis approaches. Each model was fit using either a Cox PH model (**CP** approach) or a Stratified Cox (SC) PH model (**Stratified CP, Gap Time, Marginal** approaches) that controlled for the covariate **smoking**.

Table T.1. Comparison of Results for the Treatment Variable (**tx**) Obtained from No-Interaction Models[a] Across Four Methods (Defibrillator Study)

Model	**CP**	**Stratified CP**	**Gap Time**	**Marginal**
Parameter estimate[b]	0.0839	0.0046	−0.0018	−0.0043
Robust standard error	0.1036	0.2548	0.1775	0.2579
Chi-square	0.6555	0.0003	0.0001	0.0003
p-value	0.4182	0.9856	0.9918	0.9866
Hazard ratio	1.087	1.005	0.998	0.996
95% confidence interval	(0.888, 1.332)	(0.610, 1.655)	(0.705, 1.413)	(0.601, 1.651)

[a] No-interaction SC model fitted with PROC PHREG for the **Stratified CP, Gap Time** and **Marginal** methods; no-interaction standard Cox PH model fitted for **CP** approach.
[b] Estimated coefficient of **tx** variable.

2. a. State the hazard function formula for the no-interaction model used to fit the **CP** approach.
   b. Based on the **CP** approach, what do you conclude about the effect of treatment (tx)? Explain briefly using the results in Table T.1.
   c. State the hazard function formulas for the no-interaction and interaction SC models corresponding to the use of the **Marginal** approach for fitting these data.
   d. Table T.1 gives results for "no-interaction" SC models because likelihood ratio (LR) tests comparing a "no-interaction" with an "interaction" SC model were not significant. Describe the (LR) test used for the marginal model (full and reduced models, null hypothesis, test statistic, distribution of test statistic under the null).
   e. How can you criticize the use of a no-interaction SC model for any of the SC approaches, despite the finding that the above likelihood ratio test was not significant?
   f. Based on the study description given earlier, why does it make sense to recommend the **CP** approach over the other alternative approaches?
   g. Under what circumstances/assumptions would you recommend using the **Marginal** approach instead of the **CP** approach?

Table T.2 below provides ordered failure times and corresponding risk set information that result for the 36 subjects in the above Defibrillator Study dataset using the Counting Process (**CP**) data layout format.

Table T.2. Ordered Failure Times and Risk Set Information for Defibrillator Study (CP)

Ordered failure times $t_{(f)}$	# in risk set n_f	# failed m_f	# censored in $[t_{(f)}, t_{(f+1)})$	Subject ID #s for outcomes in $[t_{(f)}, t_{(f+1)})$
0	36	0	0	—
33	36	2	0	6, 26
34	36	3	0	2, 7, 36
36	36	3	0	3, 10, 13
37	36	2	0	17, 28
38	36	4	0	22, 32, 33, 34
39	36	5	0	1, 8, 9, 11, 12
40	36	1	0	4
41	36	1	0	29
43	36	1	0	24
44	36	1	0	31

(*Continued on next page*)

Table T.2. (*Continued*)

Ordered failure times $t_{(f)}$	# in risk set n_f	# failed m_f	# censored in $[t_{(f)}, t_{(f+1)})$	Subject ID #s for outcomes in $[t_{(f)}, t_{(f+1)})$
45	36	2	0	5, 20
46	36	2	0	14, 25
48	36	1	0	21
49	36	1	0	35
51	36	2	0	23, 27
57	36	2	0	16, 30
58	36	2	0	18, 19
61	36	1	0	15
63	36	2	0	19, 34
64	36	3	0	13, 22, 33
65	36	2	0	2, 10
66	36	3	0	1, 6, 25
67	36	4	0	3, 7, 18, 24
68	36	2	0	5, 26
69	35	1	0	23
70	35	1	0	35
71	35	1	0	32
72	35	2	0	8, 20
73	35	1	0	29
74	35	1	0	31
76	35	1	0	17
77	35	1	0	14
78	35	2	0	11, 30
79	35	3	1	9, 15, 16, 28
80	34	2	0	4, 12
81	34	2	0	21, 36
93	34	2	0	7, 28
95	32	1	0	13
96	31	5	0	3, 6, 10, 26, 31
97	26	5	0	1, 22, 27, 33, 36
98	22	0	1	23
99	21	2	0	30, 34
100	19	1	0	2
102	18	1	0	8
105	17	1	0	32
106	16	2	0	19, 20
107	14	1	1	12, 35
108	12	1	0	11
109	11	0	2	9, 18
110	9	1	0	25
111	8	0	5	4, 14, 15, 24, 29
112	3	1	0	21
113	2	0	1	17
115	1	0	1	27

h. In Table T.2, why does the number in the risk set ($\mathbf{n}_f$) remain unchanged through failure time (i.e., day) 68, even though 50 events occur up to that time?
i. Why does the number in the risk set change from 31 to 26 when going from time 96 to 97?
j. Why is the number of failures ($\mathbf{m}_f$) equal to 3 and the number of censored subjects equal to 1 in the interval between failure times 79 and 80?
k. What 5 subjects were censored in the interval between failure times 111 and 112?
l. Describe the event history for subject #5, including his or her effect on changes in the risk set.

Based on the **CP** data layout of Table T.2, the following table (T.3) of survival probabilities has been calculated.

Table T.3. Survival Probabilities for Defibrillator Study Data Based on CP Layout

$t_{(f)}$	n_f	m_f	q_f	$S(t_{(f)}) = S(t_{(f-1)})Pr(T > t_{(f)} \mid T \geq t_{(f)})$
0	36	0	0	1.0
33	36	2	0	1 × 34/36 = .94
34	36	3	0	.94 × 33/36 = .87
36	36	3	0	.87 × 33/36 = .79
37	36	2	0	.79 × 34/36 = .75
38	36	4	0	.75 × 32/36 = .67
39	36	5	0	.67 × 31/36 = .57
40	36	1	0	.57 × 35/36 = .56
41	36	1	0	.56 × 35/36 = .54
43	36	1	0	.54 × 35/36 = .53
44	36	1	0	.53 × 35/36 = .51
45	36	2	0	.51 × 34/36 = .48
46	36	2	0	.48 × 34/36 = .46
48	36	1	0	.46 × 35/36 = .44
49	36	1	0	.44 × 35/36 = .43
51	36	2	0	.43 × 34/36 = .41
57	36	2	0	.41 × 34/36 = .39
58	36	2	0	.39 × 34/36 = .36
61	36	1	0	.36 × 35/36 = .35
63	36	2	0	.35 × 34/36 = .33
64	36	3	0	.33 × 33/36 = .31
65	36	2	0	.31 × 34/36 = .29
66	36	3	0	.29 × 33/36 = .27
67	36	4	0	.27 × 32/36 = .24
68	36	2	0	.24 × 34/36 = .22
69	35	1	0	.22 × 34/35 = .22
70	35	1	0	.22 × 34/35 = .21
71	35	1	0	.21 × 34/35 = .20
72	35	2	0	.20 × 33/35 = .19
73	35	1	0	.19 × 34/35 = .19
74	35	1	0	.19 × 34/35 = .18
76	35	1	0	.18 × 34/35 = .18
77	35	1	0	.18 × 34/35 = .17
78	35	2	0	.17 × 33/35 = .16

(*Continued on next page*)

Table T.3. (*Continued*)

$t_{(f)}$	n_f	m_f	q_f	$S(t_{(f)}) = S(t_{(f-1)})Pr(T > t_{(f)} \mid T \geq t_{(f)})$
79	35	3	1	$.16 \times 31/35 = .14$
80	34	2	0	$.14 \times 32/34 = .13$
81	34	2	0	$.13 \times 32/34 = .13$
95	32	1	0	$.13 \times 31/32 = .12$
96	31	5	0	$.12 \times 26/31 = .10$
97	26	5	0	$.10 \times 21/26 = .08$
98	22	0	1	$.08 \times 22/22 = .08$
99	21	2	0	$.08 \times 19/21 = .07$
100	19	1	0	$.07 \times 18/19 = .07$
102	18	1	0	$.07 \times 17/18 = .06$
105	17	1	0	$.06 \times 16/17 = .06$
106	16	2	0	$.06 \times 14/16 = .05$
107	14	1	1	$.05 \times 13/14 = .05$
108	12	1	0	$.05 \times 21/26 = .05$
109	11	0	2	$.05 \times 11/11 = .05$
110	9	1	0	$.05 \times 8/9 = .04$
111	8	0	5	$.04 \times 8/8 = .04$
112	3	1	0	$.04 \times 2/3 = .03$
113	2	0	1	$.03 \times 2/2 = .03$
115	1	0	1	$.03 \times 1/1 = .03$

Suppose the survival probabilities shown in Table T.3 are plotted on the y-axis versus corresponding ordered failure times on the x-axis.

m. What is being plotted by such a curve? (Circle one or more choices.)
   i. $Pr(T_1 > t)$ where T_1 = time to first event from study entry.
   ii. $Pr(T > t)$ where T = time from any event to the next recurrent event.
   iii. $Pr(T > t)$ where T = time to any event from study entry.
   iv. Pr(not failing prior to time t).
   v. None of the above.

n. Can you criticize the use of the product limit formula for $S(t_{(f)})$ in Table T.3? Explain briefly.

o. Use Table T.2 to complete the data layouts for plotting the following survival curves.

i. $S_1(t) = Pr(T_1 > t)$ where T_1 = time to first event from study entry

$t_{(f)}$	n_f	m_f	q_f	$S(t_{(f)}) = S(t_{(f-1)}) \times Pr(T_1 > t \mid T_1 \geq t)$
0	36	0	0	1.00
33	36	2	0	0.94
34	34	3	0	0.86
36	31	3	0	0.78
37	28	2	0	0.72
38	26	4	0	0.61
39	22	5	0	0.47
40	17	1	0	0.44
41	16	1	0	0.42
43	15	1	0	0.39
44	14	1	0	0.36
45	13	2	0	0.31
46	11	2	0	0.25
48	9	1	0	0.22
49	8	1	0	0.19
51	-	-	-	-
57	-	-	-	-
58	-	-	-	-
61	-	-	-	-

ii. **Gap Time** $S_{2c}(t) = Pr(T_{2c} > t)$ where T_{2c} = time to second event from first event.

$t_{(f)}$	n_f	m_f	q_f	$S(t_{(f)}) = S(t_{(f-1)}) \times Pr(T_1 > t \mid T_1 \geq t)$
0	36	0	0	1.00
5	36	1	0	0.97
9	35	1	0	0.94
18	34	2	0	0.89
20	32	1	0	0.86
21	31	2	1	0.81
23	28	1	0	0.78
24	27	1	0	0.75
25	26	1	0	0.72
26	25	2	0	0.66
27	23	2	0	0.60
28	21	1	0	0.58
29	20	1	0	0.55
30	19	1	0	0.52

(*Continued on next page*)

(Continued)

$t_{(f)}$	n_f	m_f	q_f	$S(t_{(f)}) = S(t_{(f-1)}) \times Pr(T_1 > t \mid T_1 \geq t)$
31	18	3	0	0.43
32	15	1	0	0.40
33	14	5	0	0.26
35	9	1	0	0.23
39	8	2	0	0.17
40	-	-	-	-
41	-	-	-	-
42	-	-	-	-
46	-	-	-	-
47	-	-	-	-

iii. **Marginal** $S_{2m}(t) = Pr(T_{2m} > t)$ where T_{2m} = time to second event from study entry.

$t_{(f)}$	n_f	m_f	q_f	$S(t_{(f)}) = S(t_{(f-1)}) \times Pr(T_1 > t \mid T_1 \geq t)$
0	36	0	0	1.00
63	36	2	0	0.94
64	34	3	0	0.86
65	31	2	0	0.81
66	29	3	0	0.72
67	26	4	0	0.61
68	22	2	0	0.56
69	20	1	0	0.53
70	19	1	0	0.50
71	18	1	0	0.47
72	17	2	0	0.42
73	15	1	0	0.39
74	14	1	0	0.36
76	13	1	0	0.33
77	12	1	0	0.31
78	11	2	0	0.25
79	-	-	-	-
80	-	-	-	-
81	-	-	-	-
97	-	-	-	-

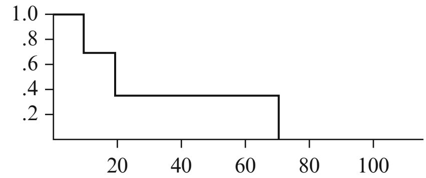
p. The survival curves corresponding to each of the data layouts (a, b, c) described in Question 14 will be different. Why?

Answers to Practice Exercises

1. T
2. F: The **Marginal** approach is appropriate if events are of different types.
3. T
4. F: The **Marginal, Stratified CP**, and **Gap Time** approaches all require a SC model, whereas the **CP** approach requires a standard PH model.
5. T
6. F: Robust estimation adjusts the **standard errors** of regression coefficients.
7. F: Robust estimation is recommended for all four approaches, not just the **CP** approach.
8. F: The P-value from robust estimation may be either larger or smaller than the corresponding P-value from nonrobust estimation.
9. F: Replace the word **Marginal** with **Stratified CP** or **Gap Time**. The Marginal approach does not use (Start, Stop) columns in its layout.
10. T
11. T
12. T
13. T
14. T
15. F: The choice among the **CP, Stratified CP, Gap Time**, and **Marginal** approaches depends on carefully considering the interpretation of each approach.
16.

$t_{(f)}$	n_f	m_f	q_f	$R(t_{(f)}\}$	$S_1(t_{(f)})$
0	3	0	0	{A, S, C}	1.00
10	3	1	0	{A, S, C}	0.67
20	2	1	0	{A, S}	0.33
70	1	1	0	{A}	0.00

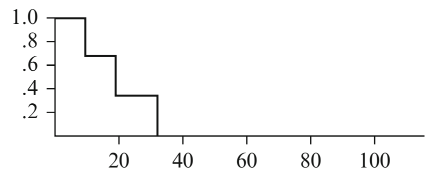
17. $\mathbf{S}_1(t)$



18.

$t_{(f)}$	n_f	m_f	q_f	$R(t_{(f)}\}$	$S_2(t_{(f)})$ Gap Time
0	3	0	0	{A, S, C}	1.00
10	3	1	0	{A, S, C}	0.67
20	2	1	0	{A, C}	0.33
30	1	1	0	{C}	0.00

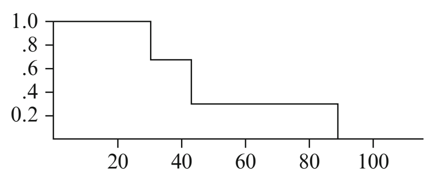
19. $\mathbf{S}_{2c}(\mathbf{t})$ **Gap Time**



20.

$t_{(f)}$	n_f	m_f	q_f	$R(t_{(f)}\}$	$S_2(t_{(f)})$ Marginal
0	3	0	0	{A, S, C}	1.00
30	3	1	0	{A, S, C}	0.67
40	2	1	0	{A, C}	0.33
90	1	1	0	{A}	0.00

21. $\mathbf{S}_{2m}(\mathbf{t})$ **Marginal**



22. All three plots differ because the risk sets for each plot are defined differently inasmuch as the failure times are different for each plot.