7
Analysis of Censored Data: Examples

7.1 Introduction

Censored survival time is the outcome of numerous studies. We select a few examples from the medical literature to give a glimpse of the scope of studies involving censored survival time. Although survival time is usually the time to death, it can be broadly referred to as the time to the occurrence of an event of interest. For example, age of onset for breast cancer can be interpreted as a survival time.

Example 7.1 Ansell et al. (1993) performed a tree-based survival analysis on 127 consecutive women with stage IIIB to stage IV ovarian cancer. Between November 1982 and July 1988, those patients had undergone surgical procedures as treatment for advanced ovarian cancer. The survival status of the patients was monitored from the time of the surgery to January 30, 1992. Eighty-four patients had died during this period of time, and the remaining 43 were still alive at the final date of the study. Hence, the survival time of the 43 alive patients was censored. The study goal is to scrutinize demographic and tumor-related prognostic (clinical, radiological, and biochemical) factors that predict survival. Based on their analysis, Ansell et al. defined three groups of patients with significantly (at the level of 0.05) different survival functions.

Example 7.2 From 1974 to 1989, 1578 patients were entered in three Radiation Therapy Oncology Group malignant glioma trials. Curran et al. (1993) used this sample to examine the associations of survival time

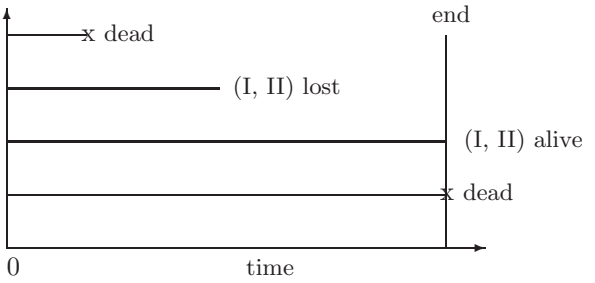
H. Zhang and B.H. Singer, *Recursive Partitioning and Applications*,
Springer Series in Statistics, DOI 10.1007/978-1-4419-6824-1_7,

to pretreatment patient status and tumor characteristics, and treatment-related indicators. The survival time was calculated from the date of the treatment to November 1991. The pretreatment factors include age, performance status, and tumor histopathology. Extent of surgery is one of the five treatment-related variables. Using the recursive partitioning technique, the authors identified six subgroups with distinct survival durations. The most important stratification is whether or not the patient was younger than 50 years of age.

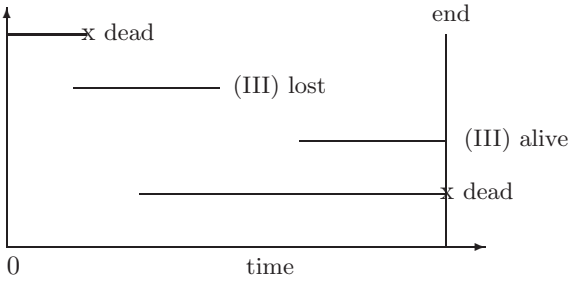
Example 7.3 The determinants of life span are complex and include genetic factors. To explore the effect of three ($H - 2^b$, $H - 2^k$, and $H - 2^d$) haplotypes on the T-cell functions and ultimately on survival, Salazar et al. (1995) conducted an experimental study using 1537 mice that were born between April 14 and July 28, 1987. The experiment ended on February 2, 1991. During the experiment period, the survival durations of 130 mice (in addition to those that were still alive at the end) were censored (not observed) because of accidental drowning of 5 and sacrifice of 125 for immunologic studies. The authors found that males lived longer than females except for $H - 2^d$ homozygotes, for which there was no sign of significant difference at the level of 0.05.

What do Examples 7.1–7.3 have in common? As in Examples 1.1–1.6, the observations from every subject include a number of predictors such as prognostic factors in Example 7.1 and genetic components in Example 7.3. What makes it more special here is the outcome of interest, which is a clearly defined, but sometimes unobservable, survival time. Depending on the nature of study, the survival time, denoted by T, may be calculated from the time of the treatment (e.g., surgery in Example 7.1) or the time of birth (e.g., Example 7.3) to the time of death (or broadly, the time when an event occurs). Due to practical constraints, we are not able to observe all subjects until after death. Thus, all studies have a clearly defined end. For example, the last day is February 2, 1991, in Example 7.3. Sometimes, the end of study may be the day when a prespecified number of study subjects have died. Those subjects that were alive at the end of the study have a censored survival time. In other words, their survival times were actually longer than what we could observe. There are also other circumstances in which we cannot observe the relevant survival time. For instance, 130 mice in Example 7.3 died from a cause other than the one of interest. They would have survived longer if they had not been killed by accidents or been sacrificed. In many human clinical trials, some subjects may be lost before the end of the study because of various health conditions or inconvenience (e.g., having moved out of the study area).

Figure 7.1 elucidates two typical study designs and three common types of censoring. In panel (a), all subjects enter into a study at the same time. When the study ends on a planned date, type I censoring occurs. If the



(a) All subjects enter into the study at the same time



(b) Subjects enter into the study at the different times

FIGURE 7.1. Three types of censoring

study is terminated after a given number of subjects have died, we have type II censoring. For both types of censoring, subjects may be either alive at the end of study or lost to follow-up during the study period. In panel (b), subjects enter into a study at different times. The censoring is classified as type III. In almost all cases where the tree-based methods have been applied, types I and III, called random censoring, are involved. Type II censoring times among subjects are not independent. We will not discuss further the distinction between random and nonrandom censoring. All of these types are *right* censoring, or censoring to the right.

Although we do not pursue it here, left censoring and interval (double) censoring also arise in practice (e.g., Peto 1973). Particularly in AIDS (acquired immunodeficiency syndrome) research, estimating the time from the HIV (human immunodeficiency virus) infection to the development of AIDS, called the incubation period, is very important to the control and prevention of AIDS (e.g., Brookmeyer 1991). The difficulty is that the time of HIV infection is unknown and the incubation period is left-censored. Supposing that the duration from the onset of HIV infection to the AIDS death is of interest, interval censoring occurs.

In summary, we cannot observe the survival time for all study subjects. To take into account the fact of survival being censored, we use δ to indicate whether a subject's survival is observed (if it is one) or censored (if it is zero). Although the survival time is the sole outcome, it involves two

response variables: the observed time, denoted by Y, and the censoring indicator. In the absence of censoring, the observed time is the survival time, and hence $Y = T$. Otherwise, the observed time is the censoring time, denoted by U. The relationship among T, Y, U, and δ is $Y = \min(T, U)$ and $\delta = I(Y = T)$, where $I(\cdot)$ is an indicator function defined as follows:

$$I(A) = \begin{cases} 1 & \text{if condition } A \text{ is met,} \\ 0 & \text{otherwise.} \end{cases} \tag{7.1}$$

We will explain later how to use the ideas expressed in Chapter 4 to analyze censored survival data. The rules of the game are essentially the same. First, a comparable "node impurity" is needed in tree growing; that is, we must define a partitioning criterion by which one node is split into two, two into more, and so on. Second, to guide tree pruning, an analogous "cost-complexity" needs to be formulated so that we can choose a "right-sized" tree, or equivalently, determine the terminal nodes. Before discussing these details, we present a tree-based survival analysis in Section 7.2 and reveal the potential of such an analysis in providing new scientific results that are not so readily attainable with other more standard methods.

7.2 Tree-Based Analysis for the Western Collaborative Group Study Data

The Western Collaborative Group Study (WCGS) is a prospective and long-term study of coronary heart disease. In 1960–61, 3154 middle-aged white males from ten large California corporations in the San Francisco Bay Area and Los Angeles entered the WCGS, and they were free of coronary heart disease and cancer. After a 33-year follow-up, 417 of 1329 deaths were due to cancer and 43 were lost to follow-up. Table 7.1 provides part of the baseline characteristics that were collected from the WCGS. A more detailed description of study design and population is available from Ragland et al. (1988). Table 7.1 gives a brief description of the predictors. In particular, body mass index (BMI) and waist-to-calf ratio (WCR) are two measures of obesity. The question of primary interest here is whether obesity as indicated by BMI and WCR is associated with the risk of cancer.

In classifying binary outcome, the impact of using different splitting criteria is relatively minor. However, the impact appears to be greater for the analysis of censored data. As we will introduce later, several criteria have been studied in the literature. We use two of them in Figure 7.2. One is based on the log-rank statistic and the other from a straightforward extension of node impurities. The next two chapters will provide in-depth discussions, but for the moment, we concentrate on the practical aspect of the analysis.

TABLE 7.1. Eight Selected Baseline Variables from the WCGS

Characteristics	Descriptive Statistics
Age	46.3 ± 5.2 years
Education	High sch. (1424), Col. (431), Grad. (1298)
Systolic blood pressure	128.6 ± 15.1 mmHg
Serum cholesterol	226.2 ± 42.9 (mg/dl)
Behavior pattern	Type A (1589), type B (1565)
Smoking habits	Yes (2439), No (715)
Body mass index	24.7 ± 2.7 (kg/m^2)
Waist-to-calf ratio	2.4 ± 0.2

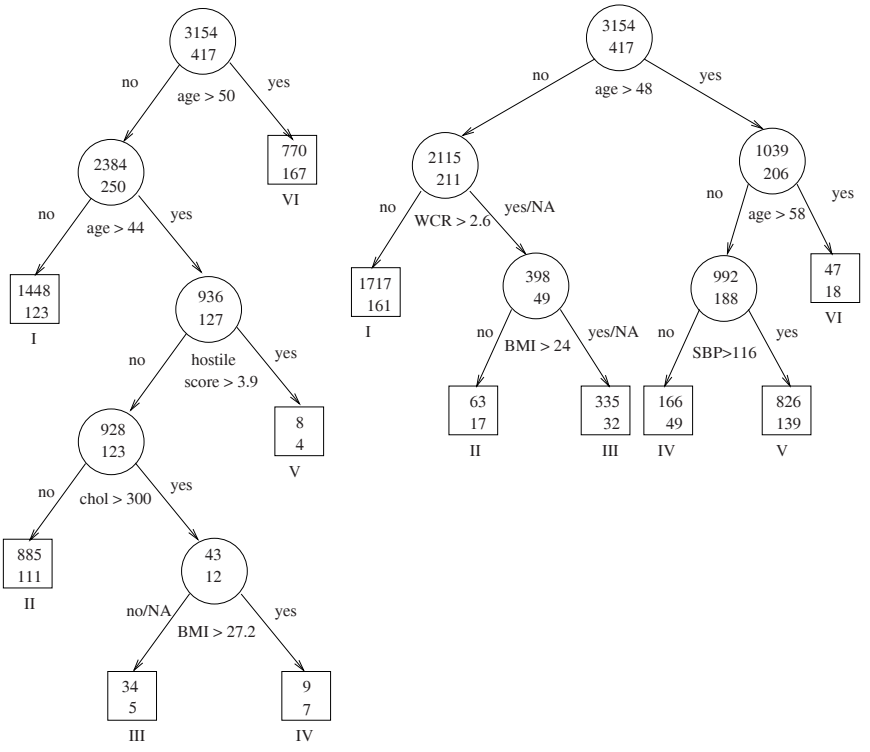


FIGURE 7.2. The survival trees using the log-rank statistic and a straightforward extension of impurity.

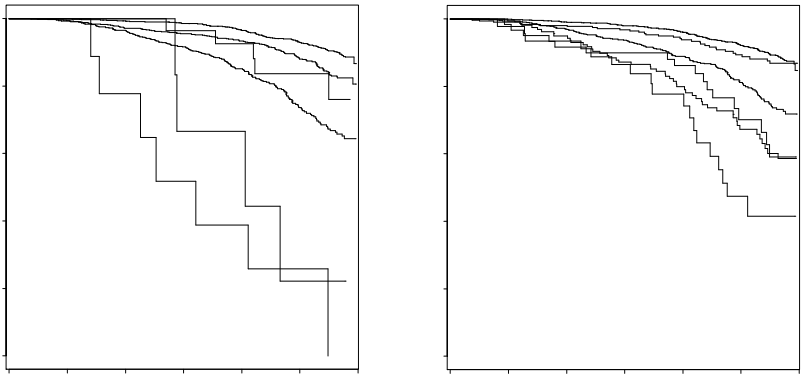


FIGURE 7.3. Kaplan–Meier curves within terminal nodes. The two panels correspond to the two trees in Figure 7.2.

How do we answer the clinical question from the survival trees? A commonly used approach is to draw Kaplan–Meier curves within all terminal nodes and then to compare these curves. Figure 7.3 is prepared following this common wisdom. Thus, the survival trees are employed as a means of stratifying the study sample. This is particularly useful when the proportionality assumption is violated in the Cox model introduced in the next chapter.

Let us first examine the tree on the left of Figure 7.2. As in the proportional hazard model (see Section 8.2.3), age and cholesterol are important attributes for survival. The hostile score seems to matter, but it requires a threshold so high (greater than 3.9) that only 8 subjects crossed the line. Instead of WCR as in the proportional hazard model, BMI, another measure of obesity, expresses some influence on the survival, but is limited to a group of 43 subjects. If we remove three survival curves for the three relatively small nodes, the left panel in Figure 7.3 suggests three major, distinct characteristics of survival, two of which are determined by age (terminal nodes I and VI). The curve for terminal node II shows that lower cholesterol levels have a dramatic protective effect on survival due to cancer.

The six survival curves on the right of Figure 7.3 display four major distinct characteristics of survival. Terminal nodes III and IV deserve our special attention. Let us point out that there are 173 missing values on BMI in terminal node III, of which 18 died from cancer. This death proportion is about the same as that among those who had BMI measured. Although subjects in terminal node I (younger and lower WCR group) had enjoyed the longest survival time, those in terminal node III had a very close survival duration. What is surprising is that this is a group with relatively high WCR and BMI. Based on the survivorship of terminal node II and the discussion above, when only one of WCR and BMI is high, the risk of death

is increased. The survivorship of terminal node V seems to raise another point. For those of age about 50 to 60, moderately high SBP is protective for survival due to cancer. These observations shed some new light on cancer death that was not uncovered from previous analyses. However, the extent of their validity warrants further investigation.