Statistics 305/605: Introduction to Biostatistical Methods for Health Sciences

Chapter 21: Survival Analysis

Jinko Graham

Introduction to Survival Analysis

- In many clinical studies, the time until the occurrence of an event is the outcome of interest.
- Examples of possible events include death, disease progression, birth, disease remission, etc.
- Events in survival analysis are defined by a transition from one discrete state to another at an instantaneous moment in time.
- Typically, not all the subjects will have experienced the event by the end of the study. This is called **censoring**.
- Censoring means that the study period ended without observing the event, either because the subject has not yet experienced it by the end of the study, or because they were lost to follow-up in some way.

Example Data:

- Data from a clinical trial to test two treatments for breast cancer (BC).
- We consider only the pre-menopausal women:
  - `time`: Time of the event (death or censoring) in the study
  - `failure`: an indicator or 0/1 variable taking values
    - 0 if censored (i.e., lost to study, or no event observed by end of study),
    - 1 if death
  - `treatment`: 1 if "treatment A" and 2 if "treatment B"
- The first few lines of data

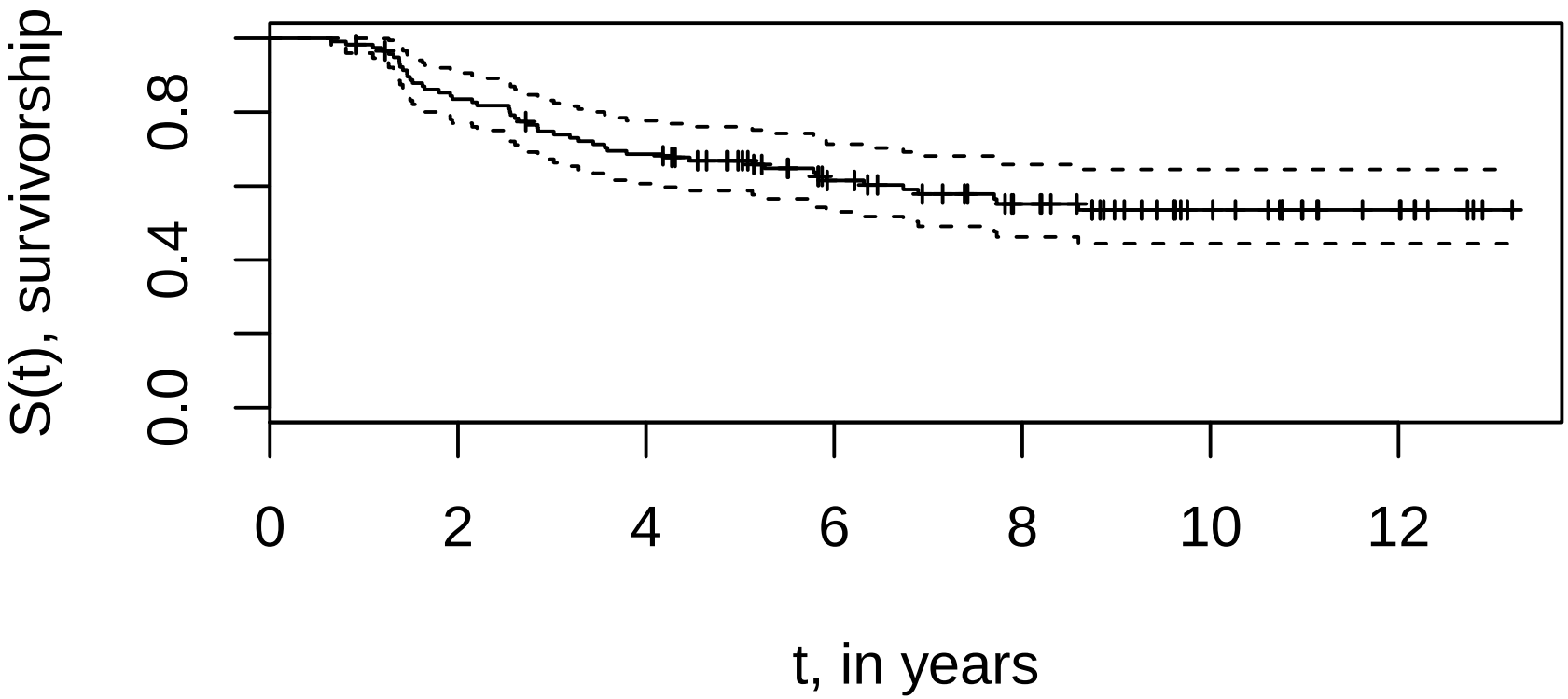
```
##        time failure treatment
## 1  6.936986       0         2
## 2  5.082192       0         1
## 3  8.745206       0         2
## 4 13.208220       0         1
## 5 12.169860       0         1
## 6  4.643836       0         2
```

Survivorship Curve

- In survival analysis, the distribution of survival times is most commonly represented by the *survivorship* curve.
- Let T denote the random survival time of a subject and t denote an observed time.
- The survivorship curve or survival function is defined as $S(t) = P(T > t)$;
  - i.e, the curve plots the chance that the survival time T is greater than some specified value $t > 0$.
  - Note that $S(0) = P(T > 0) = 1$.
- In addition to the information on $S(t)$ provided by the subjects who fail, we can use the *partial* information provided by the censored subjects.
  - E.G., A woman leaving the study after 2 years (i.e. being censored) means that she didn't die in the first 2 years.
- The next slide shows an estimate of $S(t)$ based on the BC data
  - Called the Kaplan-Meier (KM) product-limit estimate.

KM Estimate of $S(t)$

- For the BC data, the estimated survivorship function is:

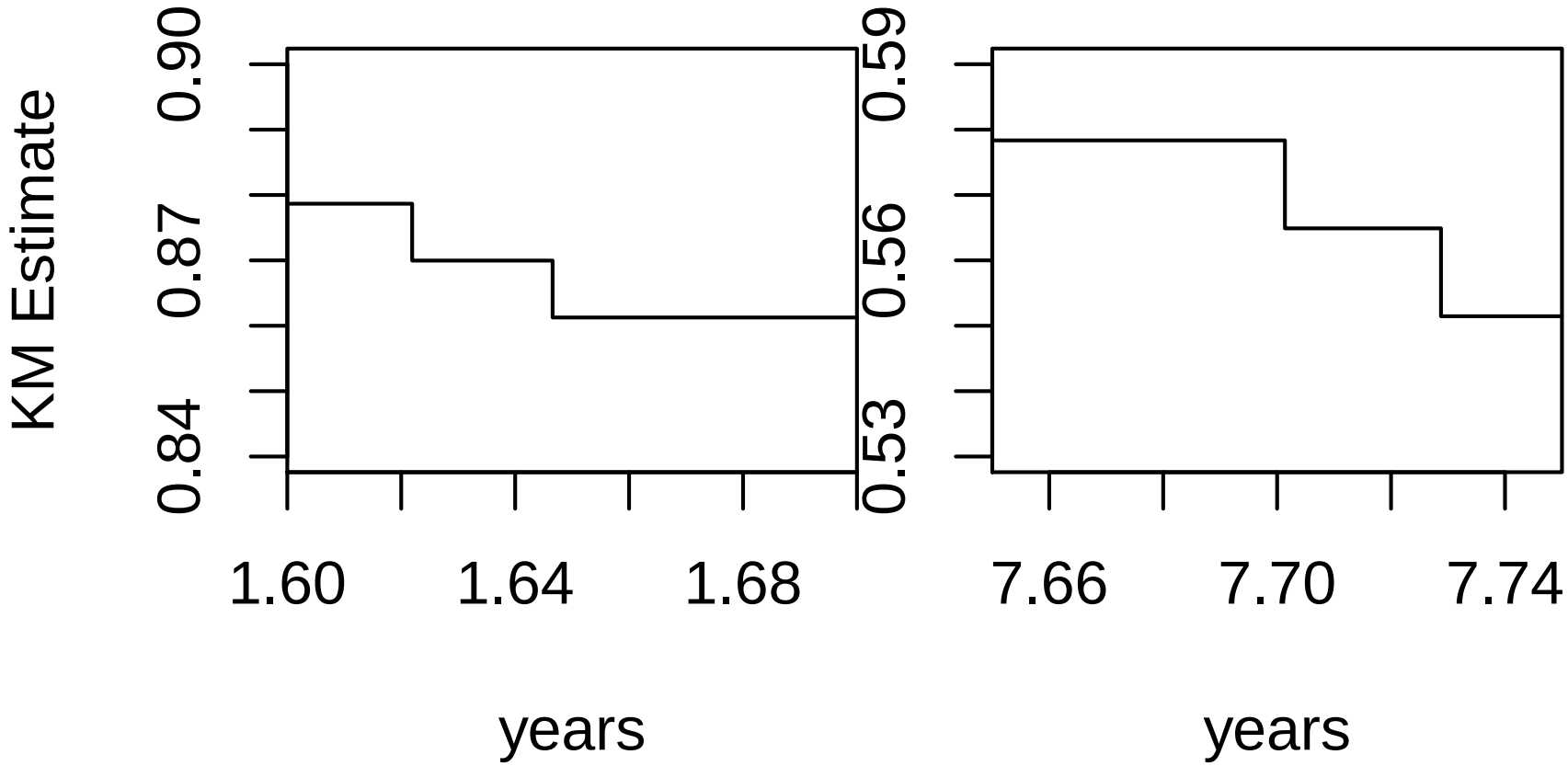


Jumps in the Estimated Survival Function

Will skip details of the derivation of $\hat{S}(t)$. Some facts:

- The KM estimate takes downward jumps at the observed failure times.
- There are no jumps in the KM estimates at censoring times.
- Instead, can think of the jump for a given censoring time as being spread out evenly amongst all the observed failure times that are greater that the censoring time.
  - In this way, late failure times stand to gain more jumps from censored subjects than earlier failure times.
  - As a result, the KM estimate takes larger drops at late failure times than at earlier failure times

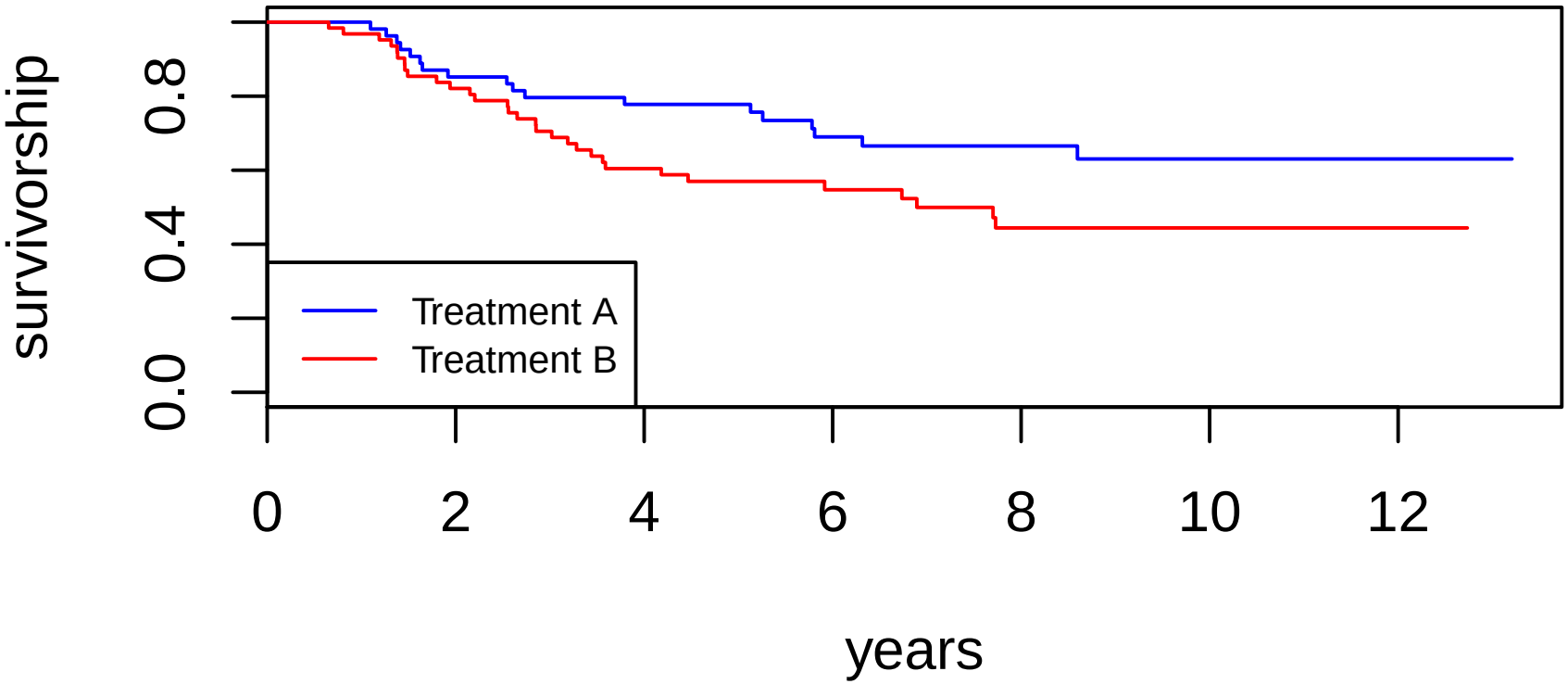
Early *vs.* Late Failure Times



- Early failure times between 1.6 and 1.7 years (left panel) and later failure times between 7.75 and 7.85 years (right panel). The y-axis is the same height (0.06) in both plots. Notice that the jumps of the KM estimate in the right panel are larger than those in the left.

KM Estimates for Subgroups

- We can stratify the analysis by explanatory variables, such as treatment.



- We estimate that the prognosis under treatment A is better than B, because the estimated treatment A curve, $\hat{S}_A$ lies above the estimated treatment B curve, $\hat{S}_B$.
  - i.e., for any time t, we estimate that the chance of surviving past t is greater for treatment A than for treatment B.

Testing for Differences in Survival: the Log-Rank Test

- To assess whether the survival curves are different, can apply the **log-rank test** of

$$H_0 : S_A(t) = S_B(t) \text{ for all } t > 0 \text{ vs. } H_a : S_A(t) \neq S_B(t).$$

- Relies on the concept of the **at-risk set** in the study at time t.
  - This is the set of all subjects who have neither failed nor been censored by time t.
- Find all failure times and corresponding at-risk sets.
- With each at-risk set, tabulate the treatment group by the failure status.

- E.G. Say there are J unique failure times in the data set.
- For each failure time, make a 2×2 table of treatment-by-failure status for individuals in the at-risk set; e.g., for the first death at $t = 0.36$:

	Failure	
	Yes	No
Treatment A	0	54
Treatment B	1	61

- Apply the **Mantel-Haenszel test** of the hypothesis that the common OR for the resulting J tables is 1 *vs.* the alternative hypothesis that the common OR is not 1.
- One way to think of the log-rank test is as a Mantel-Haenszel test of association between treatment group and failure status, that adjusts for the confounding variable failure time.

Log-rank Test for the BC Data

- The log-rank test applied to the BC data yields a statistic $X^2 = 3.8$ on 1 df, leading to a *p*value of $p = 0.0513$
  - At level $\alpha = 0.05$, there is marginal statistical evidence that the treatment group affects survival in pre-menopausal women with breast cancer.