STA 5364, Report 2.16

Carson Slater *Baylor University*

Example 4.3: Estimators of the Survival Function for Left-Truncated and Right-Censored Data

This example illustrates the application of survival function estimators for left-truncated and right-censored data using the Channing House dataset. We focus on the conditional survival function for males, estimated as:

$$S_a(t) = \prod_{a \le t_i \le t} \left(1 - \frac{d_i}{Y_i}\right), \quad t \ge a,$$

where d_i is the number of deaths and Y_i is the number at risk at time t_i.

R Code for Reproducibility

Below is the R script used to compute the conditional survival function for the Channing House dataset.

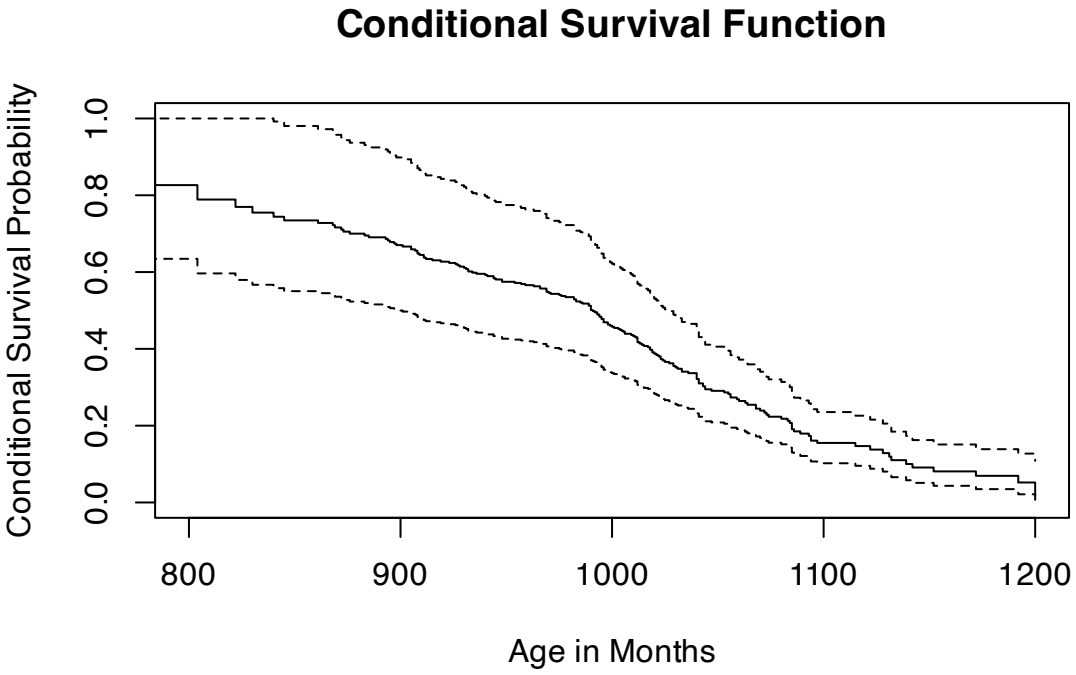
```r
# Example data for males
library("survival")
library("KMsurv")

data(channing)

# Creating the Surv object with left truncation
surv_obj <- Surv(
    channing$ageentry,
    channing$age,
    channing$death,
    type = "counting"
  )

# Kaplan-Meier estimator for left-truncated data
fit <- survfit(surv_obj ~ 1)

# Summary and plot
summary(fit)
plot(fit, conf.int = TRUE, xlab = "Age in Months",
     ylab = "Conditional Survival Probability",
     main = "Conditional Survival Function",
     xlim = c(800, 1200))
```



Explanation of Results

The survival function $S_a(t)$ is estimated for males in the dataset:

- $S_a(777) = 1.0$ (no deaths before this time),
- $S_a(781) = 1.0 \cdot \left(1 - \frac{1}{2}\right) = 0.5$,
- $S_a(800) = 0.5 \cdot \left(1 - \frac{1}{1}\right) = 0.0$.

Discussion

Figure 4.11 of Klein and Moeschberger presents the conditional survival function for males and females at ages 68 and 80 years. Using similar methods, we obtain the survival probabilities by considering only the risk set and events after the truncation age.

Conclusion

The use of conditional survival functions is crucial for left-truncated data to provide meaningful estimates. The R code presented here allows for straightforward computation and visualization of these functions.