What is Survival Analysis

Edris Safari

Bellevue University, Nebraska U.S.A.

Abstract

Exponential distributions deal with series of events occurring before or after a period of measured time. For example, are how long it takes for a car to pass through a tollgate? It is described by a CDF. Survival analysis also deals with events and time; however, it is used to study time-to-event as opposed to how long. We will introduce survival analysis and some of the related concepts in this paper.

What is Survival Analysis

Introduction

In survival analysis, the type of event is immaterial. The focus is the amount of time until the event occurs. The classical event in survival analysis is "survival" time of critically ill patients. The illness type, along with pertinent data and the length of time each patient "survived" can contribute to the analysis of the efficacy of certain medications or life changing behaviors. Given this information along with a good sampling of data, survival analysis can answer the following question:

"what is the probability that the duration or time-to-event is longer than a given time"

This question is characterized by the following function:

$$S(t) = 1 - f(t) = P(T > t)$$

Equation 1- Survival Equation

, where f(t) is the accumulation distribution function of time t, T is the duration we are looking for. $P(T > t)$ interprets to the question above.

The dataset in this case contains a record for each participant. One of the items is the time when the study ends for a participant. This end is not deterministic and can affect the analysis. The concept of "censoring" is exposed when data is missing because either the study ended before participant's record was complete (she/he passed away) or the participant stopped responding completely (left the study). Figure 1 illustrates what data is considered censored.

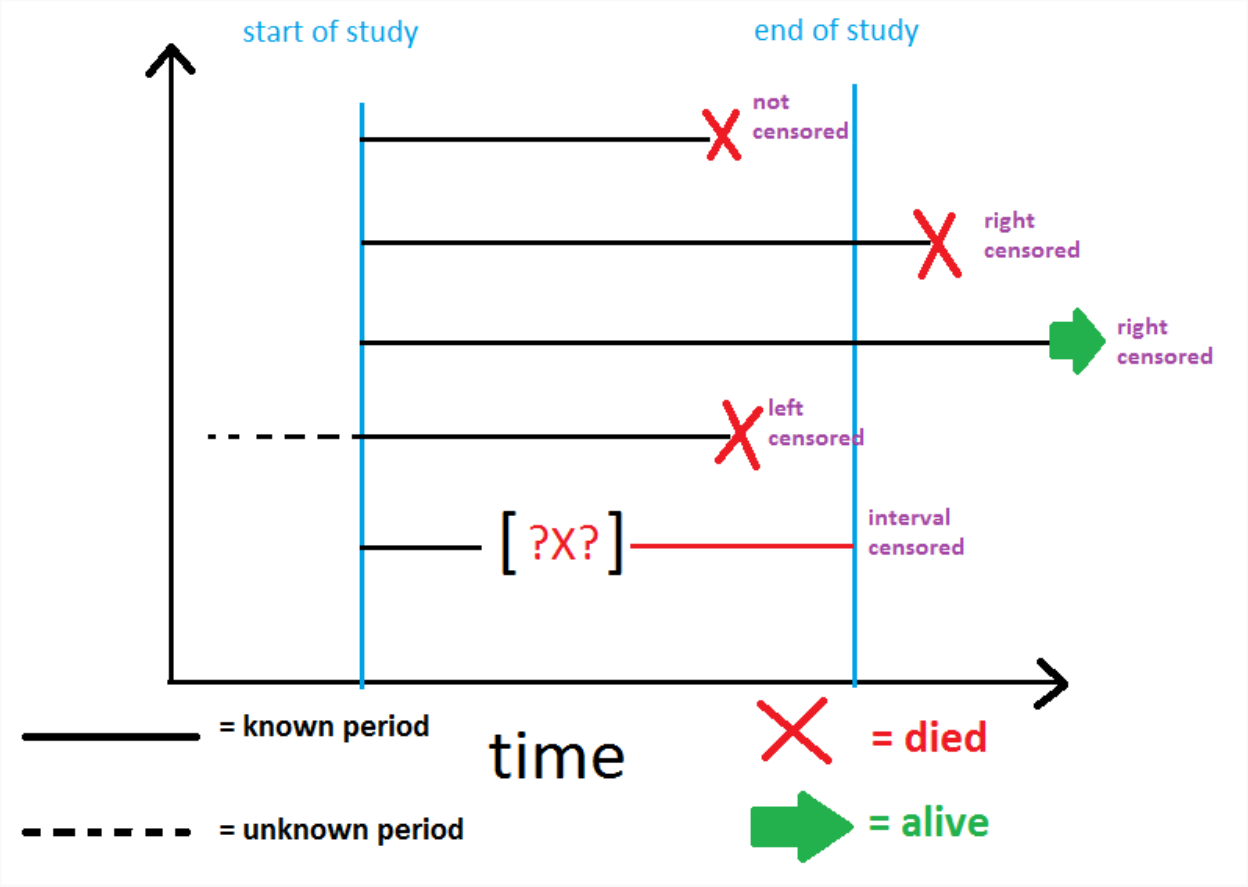

Figure 1- Censored Data

Conclusion

Survival analysis is technique in data science that helps us make smart decisions. From medical to industrial, the applications are diverse indeed. For example, in the customer retention and lifetime value, survival analysis can help business focus on the following in order to retain their customers:

- Product giveaway
- Customer refunds (by how much?)
- Marketing (how much to spend)

- Prioritization

Other applications can be found in manufacturing where lifetime of equipment in the factory are of interest. In economics and finance duration of economic events such as recession are of interest and in sociology the duration of certain societal behavior (leaning left/right/center) is of interest.

References

1. Downey, Allen B.. Think Stats: Exploratory Data Analysis . O'Reilly Media. Kindle Edition.
2. Survival Analysis - https://www.youtube.com/watch?v=iM9DitTOHBU
3. Using Survival Analysis to understand customer retention - Lorna Brightmore - https://www.youtube.com/watch?v=aKZQUaNHYb0
4. https://campus.datacamp.com/courses/survival-analysis-in-r/what-is-survival-analysis?ex=8
5. https://www.karlin.mff.cuni.cz/~pesta/NMFM404/survival.html