(IJRMST) 2020, Vol. No. 10, Jul-Dec **e-ISSN: 2455-5134, p-ISSN: 2455-9059**

Survival Analysis: Basic Concept and Censoring

Yash Vardhan Singh

ABSTRACT

In branches of medicine and biotechnology, Often until an event occurs or reoccurs, we are concerned about the time. Say, for example, the time for a hospital to collapse or the times for tumors to reappear in patients that be important. Some cancer organizations might be interested in the time until a cancer patient dies since the beginning of treatments. The latter is often referred to as survival time, whereas, the two previous examples might be termed as failure time and event time, respectively. The time for an occurrence is of concern in these cases. This time can be measured and evaluated with Survival Analysis for a variety of subjects. In recent decades the model has also been commonly used in medicine and the implementation has been thoroughly scrutinized. This article will also provide a short literary analysis on the adequacy of the application of medical journal survey analyses.

Keywords: Survival Analysis, Cancer Studies, Medical Research Review

1. SURVIVAL ANALYSIS

Survival analysis is an interesting category before an incident data analysis. Survival analysis This method is preferred and suitable for such an analysis because of several reasons. In the case of logistic regression, for example, which is mostly used in the study of the association between risk factors and disease incidence or absence, a long analysis is not feasible. This finding may lead to a downturn in the sample that would lead to concerns that we do not know about.

Survival analysis allows for the response determination of incomplete data for some subjects. For instance, observations that extend for only a certain period would result in incomplete data for some subjects who might not have developed the disease after being associated with certain risk factors. We know that they have at least some time of survival for these subjects. Of the remaining topics, we will have correct responses.

Remember that responses that are not completely recorded are censored. When no censorship is possible, standard regression procedures may be implemented, but for the following three reasons are sometimes not appropriate.

a) The responses (time to an event) is always positive and has a skewed distribution.
b) One might be rather interested in the probability of surviving past a certain point in time instead of the expected time of the event.
c) The hazardous function, that is used for regression in the survival analysis model, offers greater insight into the failure mechanism than linear regression.

Censoring occurs because

a) Some subjects, as noted earlier, may not experience the event before the study ends.
b) Some people may choose to withdraw from the study.
c) Some subjects might be lost to follow-up during the study period.

These are all examples of **right-censoring** and there are three types of them.

(IJRMST) 2020, Vol. No. 10, Jul-Dec **e-ISSN: 2455-5134, p-ISSN: 2455-9059**

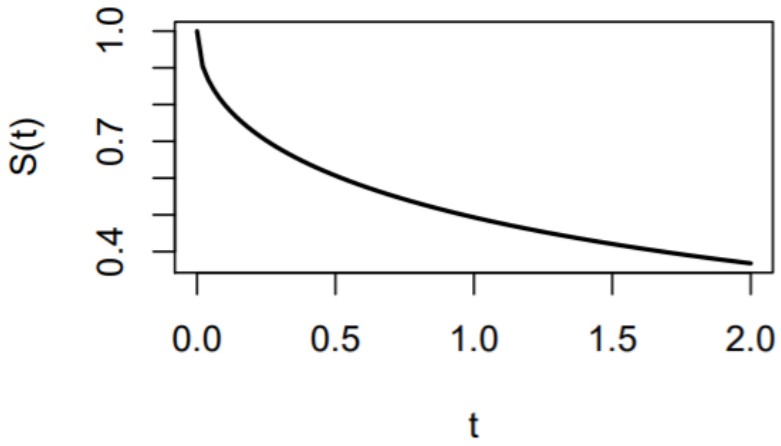
a) Example (a) where the study terminates before a subject experiences the event are examples of **fixed type I censoring.**

b) **Random type I censoring** is when all subjects are not given the same censoring time.

c) In **type II censoring,** a study ends following a pre-specified number of events.

It could be observed already that survival analysis allows for a vast array of possibilities in analyzing observations made over long periods in time. Survival analysis is a response to censoring problems as it redundancies and makes it hard to use traditional methods of graphic data discovery or presentation.

2. NOTABLE TERMINOLOGY & FORMULAE



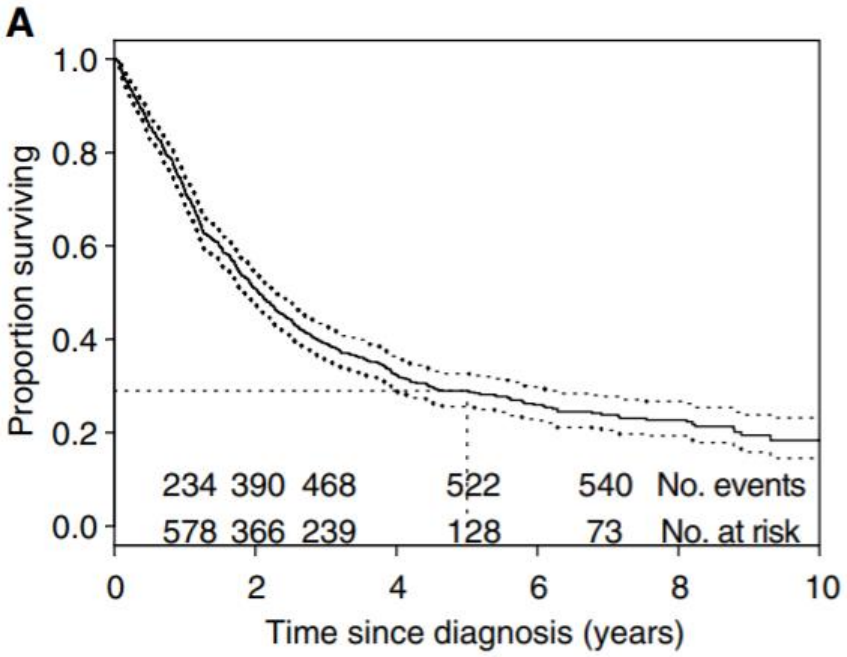
The **survival function** gives the probability of a subject surviving past time *t* is given by

$$S(t) = Pr(T > t) = 1 - F(t),$$ where T is the response variable (T > or = 0)

As *t* ranges from 0 to ∞, the function

a) is non-increasing
b) $S(0) = 1$
c) $S(t) = S(\infty) = 0$ as *t* tend ∞

Do note that the survival function is often not as smooth as it may appear to be in theory. A more representative picture would rather be the one below (based on observations of patients diagnosed with ovarian cancer).



[Credit: Clark et al. (2003)]

The **hazard function, h(t),** The instantaneous rate of interest occurring within a very narrow time frame is defined.

$$h(t) = \lim_{\Delta t \to 0} \frac{Pr(t < T \leq t + \Delta t | T > t)}{\Delta t} = \frac{f(t)}{S(t)}.$$

The **cumulative hazard** describes the accumulated risk up to time *t.*

$$H(t) = \int_0^t h(u)du.$$

If either of the functions *S(t), H(t),* or *h(t)* is known, the rest could be found using the relations below.

$$h(t) = -\frac{\partial \log(S(t))}{\partial t}$$

$$H(t) = -\log(S(t))$$

$$S(t) = \exp(-H(t))$$

3. APPLICATION OF SURVIVAL ANALYSIS MODEL IN MEDICINE

Most interest in medical studies may be the period before death. Sometimes, though, you may be worried about the time between a confirmed response and the first cancer recurrence. The use of conventional methods of regression is therefore prohibited by censored data, and thus a survival study is useful.

A review of journal articles applying the statistical model of survival analysis has led me to conclude that the model is indeed useful. The application of survey analysis has increased over the past decades, and today, it plays an important role in clinical research. The model is observed to have provided insights into patterns and has led to better and more accurate predictions. At the same time, it is also found that the statistical component of published papers is rather poor.

A rather serious problem with a majority of papers is that they are generally unable to use the statistics and present it descriptively and comprehensively. A majority of articles presented an ambiguous explanation of at least one of the endpoints of the analysis. Problems included (i) failure to evaluate the diagnosis or even prevention of non-cancer deaths as incident, and (ii) a failure to clarify whether non-relapse deaths were treated in relapse-free survival tests. It is quite astonishing and unclear if readers would be able to assess a study without a descriptive presentation. Gelber and Goldhirsch (1992) address this issue in detail. Thereupon, I also attempt to list a few key requirements that could help fix such issues by using survival analysis in clinical research.

a) **Uninformative censoring.** Even when censoring is 'noninformative' in nature are traditional approaches used to evaluate survival data with censorship. That is to say, subjects censored because of a follow-up failure would be as likely to encounter a similar event as the remaining subjects. The bias is little and the analysis of survival is useful in this way.

b) **Length of follow-up.** Time to event studies must be properly controlled to record relevant events and ensure that accurate statistical analyses are performed adequately. The proposed follow-up duration for a prospective sample would depend mainly on the extent of the illness or prognosis of the patient. As a measure of the follow-up period, the median follow-up time may be used.

c) **Completeness of follow-up.** Many subjects who do not have an incident may be considered for survival until they have been silenced, but follow-up completeness is also required. Unfair control, such as therapy arm, may distort the analysis between various groups. A predictor of inaccuracy is that a clear number of participants were missed to follow up, although they do not tell us about the time spent. A retrospective evaluation includes analysis of discrepancies in the subsequent monitoring due to the excessive dropout between the guns of the trial or several sub-groups.

d) **Cohort effect on survival.** It is generally used for treatment and other factors to be of homogeneous features during the follow-ups. However, it is likely to evolve due to innovation in ancillary treatment and/or because of a shuffling of case subjects throughout recruitment. This should be taken into consideration and adjusted in the results accordingly.

e) **Between-center differences.** Analytical methods must be consistent in each center in a multi-center trial. A diagnostic method, for example, would be close to the process diagnosis and procedures. In a case-mix analysis between centers, heterogeneity may be balanced.

4. CONCLUSION

Survey analysis is a useful statistical tool at dispense for scholars and medical practitioners to use in their studies and furtherance of medical science. Because

of censoring in observations done over a long period, the standard statistical approaches becomes redundant. Survival analysis is sensitive and prone to bias due to informative censoring and the same should be taken into adjustments. In general, the data in this model is based on two functions: the function of survival and risk. The survivor is likely to survive for some time. The survivor function The risk function refers to the absolute probability of the occurrence happening over a very short period. Comparing longevity therapies or prognostic groups also involves correction for patient-related conditions which may theoretically impact a patient's longevity time. Failure to adapt to confounders may have misleading consequences.

REFERENCES

ALTMAN, DG. STAVOLA, DL DE. LOVE, SB. STEPNIEWSKA, KA. (1995). Review of survival analyses published in cancer journals. British Journal of Cancer, 72, 511-518.

CLAR, TG. BRADBURN, MJ. LOVE, SB. ALTMAN, DG. (2003). Survival Analysis part I: Basic concepts and first analyses. British Journal of Cancer, 89, 232-238.

GELBER RD AND GOLDHIRSCH A. (1992). Reporting and interpreting adjuvant therapy in clinical trials. J. Natl Cancer Inst. Monogr., 11, 59-69.

G0TZSCHE PC. (1989). Methodology and overt and hidden bias in reports of 196 double-blind trials of nonsteroidal, anti-inflammatory drugs in rheumatoid arthritis. Controlled Clin. Trials, 10, 31-56.

CHRISTENSEN E. (1987). Multivariate survival analysis using Cox's regression model. Hepatology, 7, 1346-1358.

INTERNATIONAL COMMITTEE OF MEDICAL JOURNAL EDITORS. (1991). Uniform requirements for manuscripts submitted to biomedical journals. Br. Med. J., 302, 338-341.

NELSON W. (1969). Hazard plotting for incomplete failure data. J. Qual. Technol., 1, 27-52. PETO R, PIKE MC, ARMITAGE P, BRESLOW NE. et al. (1977). Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. Analysis and examples. Br. J. Cancer, 35, 1-39.