1.2 Basic Survival Concepts

Let $T$: Time until a specific event happen where $0 \leq T \leq \infty$. $T$ is a positive random variable.

**Cumulative Density Function (CDF)**

CDF of $T$ is defined by

$$F(t) = P(T \leq t).$$

**Properties**

- $F(0) = 0$.
- $\lim_{t \to \infty} F(t) = 1$.
- $F(t)$ is nondecreasing function. For all $t_1 \leq t_2$, $F(t_1) \leq F(t_2)$.
- $F$ is right continuous and has left limit.

**Probability Density Function (PDF)**

If $P(t \in A) = \int_{t \in A} f(t)dt$ for any subset of $\mathbb{R}$, $f$ is called the pdf of $T$. In this case, $T$ is continuous.

**Probability Mass Function (PMF)**

If $P(t \in A) = \sum_{t \in A} f(t)$ for any subset of $\mathbb{R}$, $f$ is called the pdf of $T$. In this case, $T$ is discrete.

**Survival Function**

$$S(t) = P(T \geq t).$$

**Properties**

- $S(t) = 1 - F(t)$
- $S(0) = 1$.
- $\lim_{t \to \infty} S(t) = 0$.
- $S$ is nonincreasing.
- $S$ is right continuous, has left limit.
- $f(t) = \frac{d}{dt} F(t) = -\frac{d}{dt} S(t)$.

**Hazard Function** - the instantaneous death rate

$$h(t) = \lim_{\delta t \to 0} \left\{ \frac{P(t \leq T \leq t + \delta t|T \geq t)}{\delta t} \right\}$$
Cumulative Hazard Function

\[ H(t) = \int_0^t h(u)du = -\log(S(t)) \]

Let \( T \) : Time to event and \( C \) : Drop out time.

Survival Data: \( \{T_i, C_i, i = 1, \cdots, n\} \) or \( \{X_i = \min(T_i, C_i), \delta_i = I(T_i = X_i); i = 1, \cdots, n \} \)

Some feature on Survival Analysis

1. Positive random variable

2. Censoring

In reality, an failure time of interest may not always be observed. When the follow-up observation terminates before observing the event of interest, the elapsed time from beginning the the termination is censored failure time or simply censoring time. The loss of follow up is due to end of the study, drop out from the study or death on other reasons.

Type of Censoring

- Right Censoring : An observation is known only to be greater than or equal to the observation time. \( X_i = \min(T_i, C_i) \)

  e.g. Consider a large scale animal experiment conducted at the National center for Tox-
ological Research (NCTR) in which mice were fed a particular dose of a carcinogen. The goal of the experiment was to assess the effect of the carcinogen on survival. Toward this end, mice were followed from the beginning of the experiment until death or until a pre-specified censoring time was reached, when all those still alive were sacrificed.

- **Left censoring**: Failure time is to left of the observation time. \( X_i = \max(T_i, C_i) \)
  
  e.g. In a study to determine the distribution of the time until first marijuana use among high school boys in California, and retrospective question was asked as “When did you first use marijuana?” One of the responses was “I have used it but can not recall just when the first time was.”

- **Interval censoring**: The failure time is not exactly observed, but it is observed to belong to and interval, \( T_i \in [L_i, R_i] \).
  
  e.g. In the Framingham Heart Study, the ages at which subjects first developed coronary heart disease (CHD) are usually known exactly. However, the ages of first occurrence of the subcategory angina pectoris may be known only between two clinical examinations, approximately two years.

3. **Truncation**

A condition which screens certain subjects so that the investigator will not be aware of their existence.

- **Left Truncation**: Failure time is measurable only if it is greater than a certain amount of time. (Subjects are included in the study only if they experience some event.)
  
  e.g. A study on the lifetime of automobile brake pads were conducted. A manufacturer selected a random sample of vehicles sold over the preceding 12 months at a specific group of dealers. Only cars that still had the initial tires are selected. For each car the brake pad lifetime \( t_i \) could have then been observed by following the cars prospectively.

- **Right Truncation**: Life time is measurable only if it is less than a certain amount (subjects are included in the study only if they have experienced some event).
  
  e.g. Lagakos et al. (1988) report data on the infection and induction times for patient who were infected with the AIDS virus and developed AIDS between April 1, 1978 and June 30, 1986. Only individuals who have developed AIDS prior to the end of the study
period are included in the study. Infected individuals who have yet to develop AIDS are not included in the sample.

Examples of Survival Analysis

Time to discontinuation of the use of an IUD. Following table contains the time from a woman commences use of a particular type of intrauterine device (IUD), Multiload 250, until discontinuation due to menstrual bleeding problems. $n = 18$ and all of women were between age of 18 and 35. Some of them dropped out from the study due to desire of pregnancy, no further need for a contraceptive or lost to follow up. The original study was designed to be two years. We would like to summarize the distribution of discontinuation times so that we could estimate the median time to discontinuation or the probability of a women who stop using IUD after give period of time.

0. Random variable T

1. Time origin

2. End Point

3. Censoring reasons

Prognosis for women with breast cancer. A study conducted at the Middlesex Hospital was to evaluated a histochemical marker that distinguish whether the breast cancer will metastasized to other organs in the body or not. The marker under study was a lectin known as Helix pomatia agglutinin (HPA). HPA stain positively for metastasized cells and negatively for others. A retrospective study was conducted with survival time of patients who received a simple or radical mastectomy to treat a tumor between Jan. 1969 and Dec. 1971.
study ended Jul. 1987. The objective of this study is to evaluate the evidence that those women with negative HPA staining tended to live longer after surgery than those with a positive staining. Women who were still alive at the end of study time or died with other reason was considered as censored.

0. Random variable T

1. Time origin

2. End Point

3. Censoring reasons

**Survival of multiple myeloma patients.** A study was conducted at the Medical Center of the University of West Virginia to examine the association between the values of some explanatory variables (risk factors - age, gender, levels of blood urea nitrogen, serum calcium, haemoglobin, % of plasma cells in the bone marrow and whether or not Bence-Jones protein was present in the urine) and survival time of multiple myeloma patients. \( n = 48 \) and the age of patients were between 50 and 80 years.

0. Random variable T

1. Time origin

2. End Point

3. Censoring reasons
Comparison of two treatments for prostatic cancer. A randomized, double blinded, controlled clinical trial to compare treatments (placebo vs 1.0 mg of diethylstilbestrol (DES)) for prostatic cancer was begun in 1967 by the Veteran’s Administration Cooperative Urological Research group. The objective of the study to determine the extent of any evidence that the patients with DES live longer than the other patients. Several explanatory variables (Age, Serum haemoglobin levels, size of tumor and Glaceson index) are also recorded. The survival times of patients who died from other reasons or lost of follow up are treated as censored.

0. Random variable T

1. Time origin

2. End Point

3. Censoring reasons