Hazard function

Hazard function

- Hazard function is a more popular way to describe random variables in survival analysis.
- The aim is to quantify the instantaneous risk that an event will occur at time t.
- When T is continuous, $P(T = t) = 0$ for any t.
- But we can talk about the probability that an event will occur in a small interval near time t

Hazard function

- In the definition of hazard function, we consider

$$P(t \leq T < t + \Delta t | T \geq t)$$

- However, this probability depends on Δt.
- The hazard function is defined as

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}$$

$$= \lim_{\Delta t \to 0} \frac{P(Dies\ in\ [t, t+\Delta t) | Alive\ at\ t)}{\Delta t}$$

Hazard function

- The hazard function is defined as

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}$$

- It is an instantaneous failure rate at t, given survived to t.
- Why conditional probability?
  - if individuals have already died, they are no longer at risk. We want to consider only those individuals who are still at risk at beginning of the interval.

Hazard function

I is also known as

- conditional failure rate (reliability)
- force of mortality (demography)
- intensity function (stochastic process)

Probability density function

- Recall that the probability density function

$$f(t) = \frac{dF(t)}{dt} = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t)}{\Delta t}$$

- It is the **unconditional** instantaneous failure rate at time t.

Hazard vs Density

Example 1: T denotes time from birth to death

- f(t) quantifies the risk of death at age t **for a new born**.
- h(t) quantifies the risk of death at age t **for a person of age t**.

Example 2: suppose T denotes time from surgery until recurrence for lung cancer patients

- f(t) quantifies the risk of recurrence at time t **for a patient just had surgery**.
- h(t) quantifies the risk of recurrence at time t **for a patient had surgery and hasn't had recurrent at time t**.

Why conditional?

- Interpretability: in medical studies,
  - if individuals have already died (experienced the event), they are clearly no longer at risk of the event.
  - It makes more sense to restrict to those still alive at time t.
  - Example: for a patient visits his physician, he would be more interested in conditional probabilities such as "Given that I haven't had a recurrent yet, what is my chances of having one in the next year?"
- Analytic simplification: when the data are subject to censoring, the use of hazard function simplifies analysis.

Interpretation of Hazard function

- The hazard function h(t) can be understood as the expected number of events during one unit time period after t, given $T \geq t$.
- Or 1/h(t) gives the expected length of time until event occurs, given the event has not occurred yet by time t.

Interpretation of Hazard function

- Suppose my hazard for contracting influenza at t is 0.015 with time measured in months, then over a period of one month after t, I am expect to contract influenza 0.015 times.
- If my hazard of death is 0.018 per year at time t, then I can expect to live another 1/0.018=55.5 years.