Accelerated Failure Time Model

Introduction

Example: Kidney transplant

- total of 863 Kidney transplant patients were followed until death or censoring.
- The maximum follow up time was 9.47 years.
- Patients were censored if they moved away or still alive at the end of the study.
- Goal: Investigate the effects of risk factors such as Gender, Race and Age on survival of Kidney transplant patients.

Introduction

- In linear regression model, we model the relationship between predictors to the outcome through

$$Y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_1 x_{ip} + \varepsilon_i.$$

- In generalized linear regression model, for example, logistic regression, we assume

$$\text{logit}\pi_{i1} = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_1 x_{ip}$$

Introduction

Similarly, we will focus the following two approaches to relate the outcome to predictors in survival analysis.

- Accelerated failure time model (Parametric regression)
- Cox's proportional hazard model (Semi-parametric regression)

Accelerated Failure Time (AFT) model

- Let T be the true survival time, x= $\left(x_1, \ldots, x_p\right)^T$ be explanatory variables.
- The accelerated failure time model assumes

  $Y = \log(T) = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p + \sigma\varepsilon,$

  where
  - $(\beta_0, \ldots, \beta_p)$ are regression coefficients,
  - σ i a scale parameter
  - ε is random error with mean 0 and standard deviation 1.
- It is also referred to as log-linear model.

Accelerated Failure Time (AFT) model

- The AFT model assumes that

  $$Y = \log(T) = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p + \sigma\varepsilon,$$
- It models the survival time T directly, which makes it more appealing compared with other regression methods because of this quite direct physical interpretation.
- It is often used in engineering for modeling reliability, but relatively uncommon in medical sciences.

Log-normal model

- The AFT model assumes that

$$Y = \log(T) = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p + \sigma\varepsilon.$$

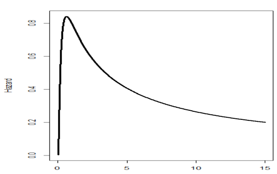
- A different distribution of ε gives different regression model.
- For example, an obvious choice is to assume that

$$\varepsilon \sim N(0,1).$$

- Then T follows a lognormal distribution.

Log-normal model

- For lognormal distribution, the hazard is of form



- May not be accurate for some real data applications.

Some popular AFT models

Distribution of ε	Distribution of T
Extreme values	Weibull ($\sigma = 1$ reduce to exponential)
Log-Gamma	Gamma
Logistic	Log-logistic
Normal	Log-normal