Statistical methods

Renjie Wei

2/19/2022

Statistical Methods

Because of censoring in survival data, the proportional hazards models are used to investigate the efficacy of a treatment (X) on a survival time T through the hazard functions. The proportional hazards model (with a single binary treatment effect) is then given by

$$h_i(t) = h_0(t)exp(x_i\beta),$$

where t is the time, $h_0(t)$ is the baseline hazard function, x_i is a binary treatment indicator variable coded 0 for control and 1 for the treatment, β is the parameter of interest, which is the log hazard ratio for the treatment effect. β measures the relative hazard reduction due to treatment in comparison to the control.

By assuming the baseline hazard function $h_0(t)$ follows one of the following form, we fit these three models on different datasets.

An exponential proportional-hazards model assumes the baseline hazard function is a constant

$$h_0(t) = \lambda$$

A Weibull proportional-hazards model assumes the baseline hazard function follows Weibull distribution, where

$$h_0(t) = \lambda\gamma t^{\gamma-1}$$

for $\gamma > 0$

A Cox proportional-hazards model leaves $h_0(t)$ unspecified.

To compare the accuracy and efficiency of the estimated treatment effects β from the three models under various baseline hazard functions, we used the following statistics: mean absolute error(MAE), variance and mean squared error(MSE)

$$MAE(\hat{\beta}) = \frac{1}{n}\sum_{i=1}^{n}|\beta - \hat{\beta}_i|$$

$$\text{var}(\hat{\beta}) = \frac{1}{n}\sum_{i=1}^{n}(\bar{\beta}_i - \hat{\beta}_i)^2$$

$$MSE(\hat{\beta}) = \frac{1}{n}\sum_{i=1}^{n}(\beta - \hat{\beta}_i)^2$$

Where i is the number of models we fitted in different size of data.