

Statistical methods

Renjie Wei

2/19/2022

Statistical Methods

Because of censoring in survival data, in our study, 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 we fit the following three models with simulated data generated by using different distribution of time.

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.

We also used Gompertz distribution and its mix with exponential distribution as baseline hazard functions inorder to evaluate the robustness of these three models, characteristics of these distributions is shown by the table below

Table 1: Characteristic of Exponential, Weibull and Gompertz distributions

	Exponential	Weibull	Gompertz
Parameter	$\lambda > 0$	$\lambda > 0, \gamma > 0$	$\lambda > 0, \alpha \in (-\infty, \infty)$
Hazard function	$h_0(t) = \lambda$	$h_0(t) = \lambda\gamma t^{\gamma-1}$	$h_0(t) = \lambda \exp(\alpha t)$
Cumulative hazard function	$H_0(t) = \lambda t$	$H_0(t) = \lambda t^\gamma$	$H_0(t) = \frac{\lambda}{\alpha}(\exp(\alpha t) - 1)$
Density function	$f_0(t) = \lambda \exp(-\lambda t)$	$f_0(t) = \lambda\gamma t^{\gamma-1} \exp(-\lambda t^\gamma)$	$f_0(t) = \lambda \exp(\alpha t) \exp(\frac{\lambda}{\alpha}(1 - \exp(\alpha t)))$
Survival function	$S_0(t) = \exp(-\lambda t)$	$S_0(t) = \exp(-\lambda t^\gamma)$	$f_0(t) = \exp(\frac{\lambda}{\alpha}(1 - \exp(\alpha t)))$

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.