A Study of the Robustness of Cox’s Proportional Hazards Model Used in Testing for Covariate Effects

Proportional hazards (PH) models are a class of semi-parametric survival models. Accelerated failure time (AFT) models are parametric models that provide alternatives to the commonly used proportional hazards models. Both models relate survival time to covariates. A proportional hazards model assumes that the effect of a covariate is to multiply the hazard by some constant. An AFT model assumes that the effect of a covariate is to multiply event time by a constant.

In order to study and compare the performance of these models, we applied PH analysis to data simulated using AFT models. Accelerated lifetime model data were simulated with one and three covariates as well as factors: sample sizes (ranging from 10 to 50), censoring rates (ranging from 0.2 to 0.9) and three distributions (Weibull, Lognormal, Loglogistic). Each combination of the factors was replicated 1000 times in order to estimate power and type I error rates. I also carried out a goodness-of-fit tests on simulated data to make sure they follow the distributions specified.

When sample size is small ($n \leq 20$) and censoring rate is high ($p \geq 0.8$), both analyses have problems of high nonconvergence rates (NR) and type I error rates. In this case, PH and AFT analyses are not suitable for hypothesis testing. But when sample size is 20 or above and censoring rate is 0.8 or below, both analyses perform well and AFT analysis has better NR than PH. PH analysis is quite stable and robust for small sample sizes and high censoring rates with respect to NR and type I error rates. I used McNemar's tests to compare PH and AFT analyses in terms of type I error rate difference (ERD), and maximum power difference (MPD). The ERD test shows that Type I error rates don't differ when sample size is big ($n \geq 50$) and censoring rate is low ($p \leq 0.3$). But the MPD test shows that the AFT analysis has a slightly higher power than PH analysis. Overall, we concluded that PH analysis is robust in hypothesis testing for covariate effects using data generated from an AFT model.