

NONPARAMETRIC MODELLING OF CLUSTERED SURVIVAL DATA

by

Iris Ivy M. Gauran

A thesis submitted in partial fulfillment
of the requirements for the degree of

Master of Science (Statistics)

School of Statistics
University of the Philippines
Diliman, Quezon City

March 2012

ABSTRACT

We postulate a nonparametric regression model to characterize clustered survival data. The postulated model incorporates the random clustering effect into the Cox Proportional Hazards model. The model was subsequently estimated via the backfitting algorithm.

The simulation study yield evidence that clustered survival data can be better characterized in a nonparametric model. Predictive accuracy in the nonparametric model increases with the number of clusters. The distribution of the random component intended to account for clustering effect also contributes into the model. As the functional form of the covariate departs from linearity, the nonparametric model is becoming more advantageous over the parametric method. Furthermore, the nonparametric model is better than the parametric model when the data is highly heterogeneous and/or there is misspecification error.

Keywords: *Survival Analysis; Clustered Data; Nonparametric Regression; Backfitting Algorithm; Random Effects; Generalized Additive Models*