ANALYSIS OF SURVIVAL FUNCTIONS IN PREDICTING LENGTH OF STAY IN FLORIDA HOSPITALS

WEBSTER BENJAMIN\textsuperscript{1} AND SEN PALI\textsuperscript{1}\*  
\textsuperscript{1}University of North Florida, Jacksonville Florida, 32224, USA.

AUTHORS' CONTRIBUTIONS  
This work was carried out in collaboration between both authors. Author SP designed the study, wrote the protocol and interpreted the data. Author WB anchored the field study, gathered the initial data and performed preliminary data analysis. Both authors searched the literature and produced the initial draft and approved the final manuscript.

Received: 10\textsuperscript{th} November 2016  
Accepted: 21\textsuperscript{st} December 2016  
Published: 23\textsuperscript{rd} December 2016

ABSTRACT

Statistical methodology and data analytics have avenues of exploring relationships among observed variables that are qualitative and quantitative in nature. The main objective of this study is to show that there is not a single “best” model to predict the length of stay of elderly patients; but rather that there is a preferred model for different age groups with various health conditions. We investigate a large amount of public data that are collected for the Agency for Health Care Administration and suggest possible predictive models to interpret its outcomes. Our data consist of every Medicare inpatient hospital discharge record related in the state of Florida 2011 related to the following primary diagnoses: Acute Myocardial Infarction, Heart Failure, and Pneumonia. The response variable is duration of stay in days. The nature of the predictor variables is either categorical or ordinal. We use an Accelerated Failure Time model and a Cox Proportional Hazard model for the right-censored response time in order to analyze related distribution functions. We interpret the effect of sex, primary diagnosis, age, inclusion of respiratory charges, and severity of illness as explanatory variables and use these to rank the patients in terms of expected length of stay. We use an extensive amount of visual display to substantiate the outcomes. The result includes expected instantaneous rate of change on the hazard functions of Accelerated Failure Time and Cox Proportional Hazard models, as well as the Kaplan Meier estimates. The study results indicate the importance of using multiple model types when analyzing any data which incorporates failure time data.

Keywords: Accelerated failure time model; Cox proportional hazard model; Kaplan Meier estimates; saturated model; reduced model; survival analysis.

*Corresponding author: Email: psen@unf.edu;