Survival Analysis: Models and Applications

Liu, X


Table of Contents

1 Introduction
  1.1 What is survival analysis and how is it applied
  1.2 The history of survival analysis and its progress
  1.3 General features of survival data structures
  1.4 Censoring
  1.5 Time scale and the origin of time
  1.6 Basic lifetime functions
  1.7 Organization of the book and data used for illustrations
  1.8 Criteria for performing survival analysis
2 Descriptive Approaches of Survival Analysis
  2.1 The Kaplan-Meier (product-limit) the Nelson-Aalen estimators
    2.1.1 Kaplan-Meier estimating procedures with or without censoring
    2.1.2 Formulation of the Kaplan-Meier and Nelson-Aalen estimators
    2.1.3 Variance and standard error of the survival function
    2.1.4 Confidence intervals and confidence bands of the survival function
  2.2 The Life table method
    2.2.1 Life table indicators
    2.2.2 Multistate life tables
    2.2.3 Illustration: Life table estimates for older Americans
  2.3 Group comparison of survival curves
    2.3.1 Logrank test for survival curves of two groups
    2.3.2 The Wilcoxon rank sum test on survival curves of two groups
    2.3.3 Comparison of survival curves for more than two groups
    2.3.4 Illustration: Comparison of survival curves between married and unmarried persons
  2.4 Summary
3 Basic Concepts and Specifications of Survival Analysis
  3.1 Exponential distribution
  3.2 Weibull distribution and extreme value theory
  3.3 Gamma distribution
  3.4 Lognormal distribution
  3.5 Log-logistic distribution
  3.6 Gompertz distribution and Gompertz-typed hazard models
  3.7 Hypergeometric distribution
3.8 Other distributions

3.9 Summary

4 Parametric Regression Models of Survival Analysis

4.1 General specifications and inferences of parametric regression models

4.1.1 Specifications of parametric regression models on hazard function

4.1.2 Specifications of accelerated failure time regression models

4.1.3 Inferences of parametric regression models and likelihood functions

4.2 Exponential regression models

4.2.1 Exponential regression model on the hazard function

4.2.2 Exponential accelerated failure time regression model

4.2.3 Illustration: Exponential regression model on marital status and survival among older Americans

4.3 Weibull regression models

4.3.1 Weibull hazard regression model

4.3.2 Weibull accelerated failure time regression model

4.3.3 Conversion of Weibull proportional hazard and AFT parameters

4.3.4 Illustration: A Weibull regression model on marital status and survival among older Americans

4.4 Log-logistic regression models

4.4.1 Specifications of Log-logistic AFT regression model

4.4.2 Retransformation of AFT parameters to untransformed Log-logistic parameters

4.4.3 Illustration: The log-logistic regression model on marital status and survival among the oldest old Americans

4.5 Other parametric regression models

4.5.1 The lognormal regression model

4.5.2 Gamma distributed regression models

4.6 Parametric regression models with interval censoring

4.6.1 Inference of parametric regression models with interval censoring

4.6.2 Illustration: A parametric regression model with independent interval censoring

4.7 Summary

5 The Cox Proportional Hazard Regression Model and Advances

5.1 The Cox semi-parametric regression model

5.1.1 Basic specifications of the Cox proportional hazard model

5.1.2 Partial likelihood

5.1.3 Procedures of maximization and hypothesis testing on partial likelihood

5.2 Estimation of the Cox hazard model with tied survival times

5.2.1 Discrete-time logistic regression model

5.2.2 Approximate methods handling ties in proportional hazard model

5.2.3 Illustration on tied survival data: Smoking cigarettes and the mortality of older Americans

5.3 Estimation of survival functions from the Cox proportional hazard model

5.3.1 The Kalbfleisch-Prentice method

5.3.2 The Breslow method
5.3.3 Illustration: Comparing survival curves for smokers and nonsmokers among older Americans

5.4 The hazard rate model with time-dependent covariates
5.4.1 Specifications of time-dependent covariates
5.4.2 The hazard rate model with time-dependent covariates
5.4.3 Illustration: A hazard model on time-dependent marital status and the mortality among older Americans

5.5 Stratified proportional hazard rate model
5.5.1 Specifications of the stratified hazard rate model
5.5.2 Illustration: Smoking cigarettes and the mortality of older Americans with stratification on three age groups

5.6 Left truncation, Left censoring, and interval censoring
5.6.1 Cox models with left truncation, left censoring, or interval censoring
5.6.2 Illustration: Analyzing left truncated survival data on smoking cigarettes and the mortality of unmarried older Americans

5.7 Qualitative factors and local tests
5.7.1 Qualitative factors and their scaling approaches
5.7.2 Local tests
5.7.3 Illustration of local tests: Educational attainment and the mortality of older Americans

5.8 Summary

6 Counting Processes and Diagnostics of the Cox Model
6.1 Counting processes and the martingale theory
6.1.1 Counting processes
6.1.2 The martingale theory
6.1.3 Stochastic integrated processes as martingale transforms
6.1.4 Martingale central limit theorems
6.1.5 Counting process formulation for the Cox model
6.2 Residuals of the Cox proportional hazard model
6.2.1 Cox-Snell residuals
6.2.2 Schonfeld residuals
6.2.3 Martingale residuals
6.2.4 Score residuals
6.2.5 Deviance residuals
6.2.6 Illustration: Residual analysis on the Cox model of smoking cigarettes and the mortality of older Americans

6.3 Assessment of proportional hazards assumption
6.3.1 Checking proportionality by adding a time-dependent variable
6.3.2 Andersen’s plots for checking proportionality
6.3.3 Check of proportionality with Schonfeld residuals
6.3.4 The Arjas plots
6.3.5 Checking proportionality with cumulative sums of martingale-based residuals
6.3.6 Illustration: Checking proportionality assumption in the Cox model for the effect of age on the mortality of older Americans

6.4 Checking the functional form of a covariate
6.4.1 Checking the functional form of a covariate using different link functions
6.4.2 Checking the function form with cumulative sums of martingale-based residuals

6.4.3 Illustration: Checking the functional form of age in the Cox model on the mortality of older Americans

6.5 Identification of influential observations

6.5.1 The likelihood displacement statistic approximation

6.5.2 LMAX statistic for identification of influential observations

6.5.3 Illustration: Checking influential observations in the Cox model on the mortality of older Americans

6.6 Summary

7 Competing Risks Models and Repeated Events

7.1 Competing risks hazard rate models

7.1.1 Latent failure times of competing risks and model specifications

7.1.2 Competing risks models and the likelihood function without covariates

7.1.3 Inferences for competing risks models with covariates

7.1.4 Competing risks model using the multinomial logit regression

7.1.5 Competing risks model with dependent failure types

7.1.6 Illustration of competing risks models: Smoking cigarettes and the mortality of older Americans from three causes of death

7.2 Repeated Events

7.2.1 Andersen and Gill model (AG)

7.2.2 PWP total-time and gap-time models (PWP-CP, PWP-GT)

7.2.3 The WLW model and extensions

7.2.4 Proportional rate and mean functions of repeated events

7.2.5 Illustration: The effects of a medical treatment on repeated patient visits

7.3 Summary

8 Structural Hazard Rate Regression Models

8.1 Some thoughts about the structural hazard regression models

8.2 Structural hazard rate model with error retransformation

8.2.1 Model specification

8.2.2 The estimation of the full model

8.2.3 The estimation of reduced-form equations

8.2.4 Decomposition of causal effects on hazard rates and survival functions

8.2.5 Illustration: The effects of veteran status on the mortality of older Americans and its pathways

8.3 Summary

9 Special Topics

9.1 Informative Censoring

9.1.1 Selection model

9.1.2 Sensitivity analysis models

9.1.3 Comments on current models handling informative censoring

9.2 Bivariate and multivariate survival functions

9.2.1 Inferences of bivariate survival model
9.2.2 Estimation of bivariate and multivariate survival models
9.2.3 Illustration of marginal models handling multivariate survival data
9.3 Frailty models
9.3.1 Hazard models with individual frailty
9.3.2 The correlated frailty model
9.3.3 Illustration of frailty models: The effect of veteran status on the mortality of older Americans revisited
9.4 Mortality crossovers and the maximum lifespan
9.4.1 Basic specifications
9.4.2 Relative acceleration of hazard rate and time of mortality crossing
9.4.3 Mathematical conditions for maximum lifespan and mortality crossover
9.5 Survival convergence and the preceding mortality crossover
9.5.1 Mathematical proofs for survival convergence and mortality crossovers
9.5.2 Simulations
9.5.3 Explanations for survival convergence and the preceding mortality crossover
9.6 Sample size required and power analysis
9.6.1 Calculation of sample size required
9.6.2 Illustration: Calculating sample size required
9.7 Summary
Appendix A The Delta Method
Appendix B. Approximation of the Variance-covariance Matrix for the Predicted Probabilities from Results of the Multinomial Logit Model
Appendix C. Simulated Patient Data on Treatment of PTSD (n = 255)
Appendix D. SAS Code for Derivation of \( \varphi \) estimates in Reduced-form Equations
Appendix E. The Analytic Results of References
Index