

# APPLIED SURVIVAL ANALYSIS

Regression Modeling of Time-to-Event Data

Second Edition

DAVID W. HOSMER

University of Massachusetts  
School of Public Health and Health Sciences  
Department of Public Health  
Division of Biostatistics and Epidemiology  
Amherst, MA

STANLEY LEMESHOW

The Ohio State University  
College of Public Health  
Center for Biostatistics  
Columbus, OH

SUSANNE MAY

University of California, San Diego  
Department of Family & Preventative Medicine  
Division of Biostatistics and Bioinformatics  
La Jolla, CA



A John Wiley & Sons, Inc., Publication

# Contents

Preface	xi
1 Introduction to Regression Modeling of Survival Data	1
1.1 Introduction, 1	
1.2 Typical Censoring Mechanisms, 3	
1.3 Example Data Sets, 9	
Exercises, 13	
2 Descriptive Methods for Survival Data	16
2.1 Introduction, 16	
2.2 Estimating the Survival Function, 17	
2.3 Using the Estimated Survival Function, 27	
2.4 Comparison of Survival Functions, 44	
2.5 Other Functions of Survival Time and Their Estimators, 59	
Exercises, 65	
3. Regression Models for Survival Data	67
3.1 Introduction, 67	
3.2 Semi-Parametric Regression Models, 69	
3.3 Fitting the Proportional Hazards Regression Model, 72	
3.4 Fitting the Proportional Hazards Model with Tied Survival Times, 85	
3.5 Estimating the Survival Function of the Proportional Hazards Regression Model, 87	
Exercises, 90	

4. Interpretation of a Fitted Proportional Hazards Regression Model	92
4.1 Introduction, 92	
4.2 Nominal Scale Covariate, 94	
4.3 Continuous Scale Covariate, 106	
4.4 Multiple-Covariate Models, 108	
4.5 Interpreting and Using the Estimated Covariate-Adjusted Survival Function, 121	
Exercises, 130	
5. Model Development	132
5.1 Introduction, 132	
5.2 Purposeful Selection of Covariates, 133	
5.2.1 Methods to examine the scale of continuous covariates in the log hazard, 136	
5.2.2 An example of purposeful selection of covariates, 141	
5.3 Stepwise, Best-Subsets and Multivariable Fractional Polynomial Methods of Selecting Covariates, 153	
5.3.1 Stepwise selection of covariates, 154	
5.3.2 Best subsets selection of covariates, 159	
5.3.3 Selecting covariates and checking their scale using multivariable fractional polynomials, 162	
5.4 Numerical Problems, 166	
Exercises, 168	
6. Assessment of Model Adequacy	169
6.1 Introduction, 169	
6.2 Residuals, 170	
6.3 Assessing the Proportional Hazards Assumption, 177	
6.4 Identification of Influential and Poorly Fit Subjects, 184	
6.5 Assessing Overall Goodness-of-Fit, 191	
6.6 Interpreting and Presenting Results From the Final Model, 195	
Exercises, 205	
7. Extensions of the Proportional Hazards Model	207
7.1 Introduction, 207	
7.2 The Stratified Proportional Hazards Model, 208	
7.3 Time-Varying Covariates, 213	
7.4 Truncated, Left Censored and Interval Censored Data, 228	
Exercises, 241	

8. Parametric Regression Models	244
8.1 Introduction, 244	
8.2 The Exponential Regression Model, 246	
8.3 The Weibull Regression Model, 260	
8.4 The Log-Logistic Regression Model, 273	
8.5 Other Parametric Regression Models, 283	
Exercises, 283	
<b>9. Other Models and Topics</b>	<b>286</b>
9.1 Introduction, 286	
9.2 Recurrent Event Models, 287	
9.3 Frailty Models, 296	
9.4 Nested Case-Control Studies, 308	
9.5 Additive Models, 314	
9.6 Competing Risk Models, 329	
9.7 Sample Size and Power, 340	
9.8 Missing Data, 346	
Exercises, 351	
 Appendix 1   The Delta Method	 355
 Appendix 2   An Introduction to the Counting Process Approach to Survival Analysis	  359
 Appendix 3   Percentiles for Computation of the Hall and Wellner Confidence Band	  364
 References	 365
 Index	 383