Contents

Preface v

Notation 1

1 Introduction 8

1.1 Motivation, Bibliographic History, and an Overview of the book. 8

1.2 Tour through the General Estimation Problem. 16

1.2.1 Estimation in a high-dimensional full data model 17

1.2.2 The curse of dimensionality in the full data model 21

1.2.3 Coarsening at random 23

1.2.4 The curse of dimensionality revisited 27

1.2.5 The observed data model 40

1.2.6 General method for construction of locally efficient estimators 40

1.2.7 Comparison with maximum likelihood estimation 45

1.3 Example: Causal Effect of Air Pollution on Short-Term Asthma Response 48

1.4 Estimating Functions 55

1.4.1 Orthogonal complement of a nuisance tangent space 55

1.4.2 Review of efficiency theory 61

1.4.3 Estimating functions 62

1.4.4 Orthogonal complement of a nuisance tangent space in an observed data model 64
1.4.5 Basic useful results to compute projections 68
1.5 Robustness of Estimating Functions 69
  1.5.1 Robustness of estimating functions against misspecification of linear convex nuisance parameters 69
  1.5.2 Double robustness of observed data estimating functions 77
  1.5.3 Understanding double robustness for a general semiparametric model 79
1.6 Doubly robust estimation in censored data models 81
1.7 Using Cross-Validation to Select Nuisance Parameter Models 93
  1.7.1 A semiparametric model selection criterion 94
  1.7.2 Forward/backward selection of a nuisance parameter model based on cross-validation with respect to the parameter of interest 97
  1.7.3 Data analysis example: Estimating the causal relationship between boiled water use and diarrhea in HIV-positive men 99

2 General Methodology 102
  2.1 The General Model and Overview 102
  2.2 Full Data Estimating Functions 103
    2.2.1 Orthogonal complement of the nuisance tangent space in the multivariate generalized linear regression model (MGLM) 105
    2.2.2 Orthogonal complement of the nuisance tangent space in the multiplicative intensity model 107
    2.2.3 Linking the orthogonal complement of the nuisance tangent space to estimating functions 111
  2.3 Mapping into Observed Data Estimating Functions 114
    2.3.1 Initial mappings and reparametrizing the full data estimating functions 114
    2.3.2 Initial mapping indexed by censoring and protected nuisance parameter 124
    2.3.3 Extending a mapping for a restricted censoring model to a complete censoring model 125
    2.3.4 Inverse weighting a mapping developed for a restricted censoring model 126
    2.3.5 Beating a given RAL estimator 128
    2.3.6 Orthogonalizing an initial mapping w.r.t. G: Double robustness 131
    2.3.7 Ignoring information on the censoring mechanism improves efficiency 135
  2.4 Optimal Mapping into Observed Data Estimating Functions 137
2.4.1 The corresponding estimating equation ........................................ 139
2.4.2 Discussion of ingredients of a one-step estimator .......................... 141
2.5 Guaranteed Improvement Relative to an Initial Estimating Function .................................................. 142
2.6 Construction of Confidence Intervals .................................................. 144
2.7 Asymptotics of the One-Step Estimator .............................................. 145
  2.7.1 Asymptotics assuming consistent estimation of the censoring mechanism .................................................. 146
  2.7.2 Proof of Theorem 2.4 ................................................................. 150
  2.7.3 Asymptotics assuming that either the censoring mechanism or the full data distribution is estimated consistently .................................................. 151
  2.7.4 Proof of Theorem 2.5 ................................................................. 152
2.8 The Optimal Index ............................................................................. 153
  2.8.1 Finding the optimal estimating function among a given class of estimating functions .................................................. 159
2.9 Estimation of the Optimal Index ......................................................... 166
  2.9.1 Reparametrizing the representations of the optimal full data function .................................................. 167
  2.9.2 Estimation of the optimal full data structure estimating function .................................................. 169
2.10 Locally Efficient Estimation with Score-Operator Representation .................................................. 170

3 Monotone Censored Data ....................................................................... 172
  3.1 Data Structure and Model ................................................................. 172
  3.1.1 Cause-specific censoring ............................................................... 175
  3.2 Examples ......................................................................................... 176
    3.2.1 Right-censored data on a survival time ............................................. 176
    3.2.2 Right-censored data on quality-adjusted survival time .......................... 177
    3.2.3 Right-censored data on a survival time with reporting delay ..................... 179
    3.2.4 Univariately right-censored multivariate failure time data .......................... 181
  3.3 Inverse Probability Censoring Weighted (IPCW) Estimators ............... 183
    3.3.1 Identifiability condition ............................................................... 183
    3.3.2 Estimation of a marginal multiplicative intensity model ............................. 184
    3.3.3 Extension to proportional rate models ............................................... 191
    3.3.4 Projecting on the tangent space of the Cox proportional hazards model of the censoring mechanism .................................................. 192
  3.4 Optimal Mapping into Estimating Functions ........................................ 195
3.5 Estimation of $Q$ .................................................. 196
  3.5.1 Regression approach: Assuming that the censoring mechanism is correctly specified ............ 197
  3.5.2 Maximum likelihood estimation according to a multiplicative intensity model: Doubly robust .... 198
  3.5.3 Maximum likelihood estimation for discrete models: Doubly robust .................................. 200
  3.5.4 Regression approach: Doubly robust .......... 201
3.6 Estimation of the Optimal Index ......................... 204
  3.6.1 The multivariate generalized regression model . 205
  3.6.2 The multivariate generalized regression model when covariates are always observed .......... 206
3.7 Multivariate failure time regression model ............... 208
3.8 Simulation and data analysis for the nonparametric full data model ....................................... 211
3.9 Rigorous Analysis of a Bivariate Survival Estimate .... 217
  3.9.1 Proof of Theorem 3.2 ................................. 221
3.10 Prediction of Survival ....................................... 224
  3.10.1 General methodology ............................... 225
  3.10.2 Prediction of survival with Regression Trees ... 230
4 Cross-Sectional Data and Right-Censored Data Combined 232
  4.1 Model and General Data Structure ...................... 232
  4.2 Cause Specific Monitoring Schemes .................... 234
    4.2.1 Overview ............................................ 235
  4.3 The Optimal Mapping into Observed Data Estimating Functions .......................................... 236
    4.3.1 Identifiability condition .......................... 239
    4.3.2 Estimation of a parameter on which we have current status data .................................. 241
    4.3.3 Estimation of a parameter on which we have right-censored data .................................. 243
    4.3.4 Estimation of a joint-distribution parameter on which we have current status data and right-censored data .............................................. 244
  4.4 Estimation of the Optimal Index in the MGLM ....... 245
  4.5 Example: Current Status Data with Time-Dependent Covariates ............................................. 246
    4.5.1 Regression with current status data .............. 248
    4.5.2 Previous work and comparison with our results .. 250
    4.5.3 An initial estimator .................................. 251
    4.5.4 The locally efficient one-step estimator ........ 252
    4.5.5 Implementation issues ............................. 253
    4.5.6 Construction of confidence intervals ............. 255
6.4.1 Doubly robust estimators in marginal structural models with right-censoring .................. 334
6.4.2 Data Analysis: SPARCS .......................... 338
6.4.3 A simulation for estimators of a treatment-specific survival function ............................ 343
6.5 Structural Nested Model with Right-Censoring .................. 347
6.5.1 The orthogonal complement of a nuisance tangent space in a structural nested model without censoring 353
6.5.2 A class of estimating functions for the marginal structural nested model ......................... 357
6.5.3 Analyzing dynamic treatment regimes .................. 359
6.5.4 Simulation for dynamic regimes in point treatment studies ........................................... 360
6.6 Right-Censoring with Missingness .......................... 362
6.7 Interval Censored Data .................................. 366
6.7.1 Interval censoring and right-censoring combined 368

References 371

Author index 388

Subject index 394

Example index 397
Unified Methods for Censored Longitudinal Data and Causality
Laan, M.J. van der; Robins, J.M.
2003, XII, 399 p. 2 illus., Hardcover
ISBN: 978-0-387-95556-8