
Contents

Preface	vii
1 Introduction to Semiparametric Models	1
1.1 What Is an Infinite-Dimensional Space?	2
1.2 Examples of Semiparametric Models	3
Example 1: Restricted Moment Models	3
Example 2: Proportional Hazards Model	7
Example 3: Nonparametric Model	8
1.3 Semiparametric Estimators	8
2 Hilbert Space for Random Vectors	11
2.1 The Space of Mean-Zero q -dimensional Random Functions	11
The Dimension of the Space of Mean-Zero Random Functions	12
2.2 Hilbert Space	13
2.3 Linear Subspace of a Hilbert Space and the Projection Theorem	14
Projection Theorem for Hilbert Spaces	14
2.4 Some Simple Examples of the Application of the Projection Theorem	15
Example 1: One-Dimensional Random Functions	15
Example 2: q -dimensional Random Functions	16
2.5 Exercises for Chapter 2	19
3 The Geometry of Influence Functions	21
3.1 Super-Efficiency	24
Example Due to Hodges	24
3.2 m -Estimators (Quick Review)	29
Estimating the Asymptotic Variance of an m -Estimator	31
Proof of Theorem 3.2	34

3.3	Geometry of Influence Functions for Parametric Models	38
	Constructing Estimators	38
3.4	Efficient Influence Function	42
	Asymptotic Variance when Dimension Is Greater than One . . .	43
	Geometry of Influence Functions	45
	Deriving the Efficient Influence Function	46
3.5	Review of Notation for Parametric Models	49
3.6	Exercises for Chapter 3	50
4	Semiparametric Models	53
4.1	GEE Estimators for the Restricted Moment Model	54
	Asymptotic Properties for GEE Estimators	55
	Example: Log-linear Model	57
4.2	Parametric Submodels	59
4.3	Influence Functions for Semiparametric RAL Estimators	61
4.4	Semiparametric Nuisance Tangent Space	63
	Tangent Space for Nonparametric Models	68
	Partitioning the Hilbert Space	69
4.5	Semiparametric Restricted Moment Model	73
	The Space Λ_{2s}	77
	The Space Λ_{1s}	79
	Influence Functions and the Efficient Influence Function for the Restricted Moment Model	83
	The Efficient Influence Function	85
	A Different Representation for the Restricted Moment Model	87
	Existence of a Parametric Submodel for the Arbitrary Restricted Moment Model	91
4.6	Adaptive Semiparametric Estimators for the Restricted Moment Model	93
	Extensions of the Restricted Moment Model	97
4.7	Exercises for Chapter 4	98
5	Other Examples of Semiparametric Models	101
5.1	Location-Shift Regression Model	101
	The Nuisance Tangent Space and Its Orthogonal Complement for the Location-Shift Regression Model	103
	Semiparametric Estimators for β	106
	Efficient Score for the Location-Shift Regression Model	107
	Locally Efficient Adaptive Estimators	108
	Remarks	113
5.2	Proportional Hazards Regression Model with Censored Data	113
	The Nuisance Tangent Space	117

	The Space Λ_{2s} Associated with $\lambda_{C X}(v x)$	117
	The Space Λ_{1s} Associated with $\lambda(v)$	119
	Finding the Orthogonal Complement of the Nuisance Tangent Space	120
	Finding RAL Estimators for β	123
	Efficient Estimator	125
5.3	Estimating the Mean in a Nonparametric Model	125
5.4	Estimating Treatment Difference in a Randomized Pretest-Posttest Study or with Covariate Adjustment	126
	The Tangent Space and Its Orthogonal Complement	129
5.5	Remarks about Auxiliary Variables	133
5.6	Exercises for Chapter 5	135
6	Models and Methods for Missing Data	137
6.1	Introduction	137
6.2	Likelihood Methods	143
6.3	Imputation	144
	Remarks	145
6.4	Inverse Probability Weighted Complete-Case Estimator	146
6.5	Double Robust Estimator	147
6.6	Exercises for Chapter 6	150
7	Missing and Coarsening at Random for Semiparametric Models	151
7.1	Missing and Coarsened Data	151
	Missing Data as a Special Case of Coarsening	153
	Coarsened-Data Mechanisms	154
7.2	The Density and Likelihood of Coarsened Data	156
	Discrete Data	156
	Continuous Data	157
	Likelihood when Data Are Coarsened at Random	158
	Brief Remark on Likelihood Methods	160
	Examples of Coarsened-Data Likelihoods	161
7.3	The Geometry of Semiparametric Coarsened-Data Models	163
	The Nuisance Tangent Space Associated with the Full-Data Nuisance Parameter and Its Orthogonal Complement ..	166
7.4	Example: Restricted Moment Model with Missing Data by Design	174
	The Logistic Regression Model	179
7.5	Recap and Review of Notation	181
7.6	Exercises for Chapter 7	183

8	The Nuisance Tangent Space and Its Orthogonal Complement	185
8.1	Models for Coarsening and Missingness	185
	Two Levels of Missingness	185
	Monotone and Nonmonotone Coarsening for more than Two Levels	186
8.2	Estimating the Parameters in the Coarsening Model	188
	MLE for ψ with Two Levels of Missingness	188
	MLE for ψ with Monotone Coarsening	189
8.3	The Nuisance Tangent Space when Coarsening Probabilities Are Modeled	190
8.4	The Space Orthogonal to the Nuisance Tangent Space	192
8.5	Observed-Data Influence Functions	193
8.6	Recap and Review of Notation	195
8.7	Exercises for Chapter 8	196
9	Augmented Inverse Probability Weighted Complete-Case Estimators	199
9.1	Deriving Semiparametric Estimators for β	199
	Interesting Fact	206
	Estimating the Asymptotic Variance	206
9.2	Additional Results Regarding Monotone Coarsening	207
	The Augmentation Space Λ_2 with Monotone Coarsening	207
9.3	Censoring and Its Relationship to Monotone Coarsening	213
	Probability of a Complete Case with Censored Data	216
	The Augmentation Space, Λ_2 , with Censored Data	216
	Deriving Estimators with Censored Data	217
9.4	Recap and Review of Notation	218
9.5	Exercises for Chapter 9	220
10	Improving Efficiency and Double Robustness with Coarsened Data	221
10.1	Optimal Observed-Data Influence Function Associated with Full-Data Influence Function	221
10.2	Improving Efficiency with Two Levels of Missingness	225
	Finding the Projection onto the Augmentation Space	226
	Adaptive Estimation	227
	Algorithm for Finding Improved Estimators with Two Levels of Missingness	229
	Remarks Regarding Adaptive Estimators	230
	Estimating the Asymptotic Variance	233
	Double Robustness with Two Levels of Missingness	234

Remarks Regarding Double-Robust Estimators	236
Logistic Regression Example Revisited	236
10.3 Improving Efficiency with Monotone Coarsening	239
Finding the Projection onto the Augmentation Space	239
Adaptive Estimation	243
Double Robustness with Monotone Coarsening	248
Example with Longitudinal Data	251
10.4 Remarks Regarding Right Censoring	254
10.5 Improving Efficiency when Coarsening	
Is Nonmonotone	255
Finding the Projection onto the Augmentation Space	256
Uniqueness of $\mathcal{M}^{-1}(\cdot)$	258
Obtaining Improved Estimators with Nonmonotone	
Coarsening	261
Double Robustness	265
10.6 Recap and Review of Notation	267
10.7 Exercises for Chapter 10	270
11 Locally Efficient Estimators for Coarsened-Data	
Semiparametric Models	273
Example: Estimating the Mean with Missing Data	275
11.1 The Observed-Data Efficient Score	277
Representation 1 (Likelihood-Based)	277
Representation 2 (AIPWCC-Based)	278
Relationship between the Two Representations	278
\mathcal{M}^{-1} for Monotone Coarsening	282
\mathcal{M}^{-1} with Right Censored Data	284
11.2 Strategy for Obtaining Improved Estimators	285
Example: Restricted Moment Model with Monotone	
Coarsening	286
Some Brief Remarks Regarding Robustness	290
11.3 Concluding Thoughts	291
11.4 Recap and Review of Notation	292
11.5 Exercises for Chapter 11	293
12 Approximate Methods for Gaining Efficiency	295
12.1 Restricted Class of AIPWCC Estimators	295
12.2 Optimal Restricted (Class 1) Estimators	300
Deriving the Optimal Restricted (Class 1) AIPWCC	
Estimator	305
Estimating the Asymptotic Variance	307
12.3 Example of an Optimal Restricted	
(Class 1) Estimator	309
Modeling the Missingness Probabilities	312
12.4 Optimal Restricted (Class 2) Estimators	313

Logistic Regression Example Revisited	319
12.5 Recap and Review of Notation	321
12.6 Exercises for Chapter 12	322
13 Double-Robust Estimator of the Average Causal Treatment Effect	323
13.1 Point Exposure Studies	323
13.2 Randomization and Causality	326
13.3 Observational Studies	327
13.4 Estimating the Average Causal Treatment Effect	328
Regression Modeling	328
13.5 Coarsened-Data Semiparametric Estimators	329
Observed-Data Influence Functions	331
Double Robustness	336
13.6 Exercises for Chapter 13	337
14 Multiple Imputation: A Frequentist Perspective	339
14.1 Full- Versus Observed-Data Information Matrix	342
14.2 Multiple Imputation	344
14.3 Asymptotic Properties of the Multiple-Imputation Estimator	346
Stochastic Equicontinuity	352
14.4 Asymptotic Distribution of the Multiple-Imputation Estimator	354
14.5 Estimating the Asymptotic Variance	362
Consistent Estimator for the Asymptotic Variance	365
14.6 Proper Imputation	366
Asymptotic Distribution of $n^{1/2}(\hat{\beta}_n^* - \beta_0)$	367
Rubin's Estimator for the Asymptotic Variance	370
Summary	371
14.7 Surrogate Marker Problem Revisited	371
How Do We Sample?	373
References	375
Index	381