

bard, Nick Jewell, Susan Murphy, Dan Scharfstein and Aad van der Vaart. The enthusiasm of epidemiologists John Colford, Bill Satariano, and Ira Tager for statistical methods suitable for addressing causal questions of interest in their studies on AIDS, air pollution, drinking water, and effect of exercise on survival, has been particularly stimulating.

As always, efforts such as writing a book are a product of the community one works in. Therefore it is difficult to mention everybody who has contributed in one way or another to this book. Mark van der Laan would like to give special thanks to Martine since writing this book would have been hard without her wonderful support, and to Laura, Lars, and Robin as well for not having lost track of the important things in life. Mark van der Laan also thanks Nick Jewell for his continuous intellectual interest and moral support during the writing of this book, and Bonnie Hutchings for being an incredible asset to our Department. James Robins would like to specially thank Andrea Rotnitzky for being both intellectual soulmate and supportive friend, Susanna Kaysen for her extended hospitality at Casa Kaysen, and Valérie Ventura for just about everything.

Mark van der Laan has been supported in his research efforts over the period of work on this book (1994-2002) by a FIRST award (GM53722) grant from the National Institute of General Medical Sciences and a NIAID grant (1-R01-AI46182-01), National Institute of Health. James Robins has been supported by an NIAID grant.

Mark J. van der Laan

James M. Robins

## Contents

Preface	v
Notation	1
<b>1 Introduction</b>	<b>8</b>
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