

Table of Contents

Foreword to the Revised Edition xi

Preface xiii

1 Observational Studies and Experiments

- 1.1 Introduction 1
- 1.2 The HIP trial 4
- 1.3 Snow on cholera 6
- 1.4 Yule on the causes of poverty 9
 - Exercise set A 13
- 1.5 End notes 14

2 The Regression Line

- 2.1 Introduction 18
- 2.2 The regression line 18
- 2.3 Hooke's law 22
 - Exercise set A 23
- 2.4 Complexities 23
- 2.5 Simple vs multiple regression 26
 - Exercise set B 26
- 2.6 End notes 28

3 Matrix Algebra

- 3.1 Introduction 29
 - Exercise set A 30
- 3.2 Determinants and inverses 31
 - Exercise set B 33
- 3.3 Random vectors 35
 - Exercise set C 35
- 3.4 Positive definite matrices 36
 - Exercise set D 37
- 3.5 The normal distribution 38
 - Exercise set E 39
- 3.6 If you want a book on matrix algebra 40

4 Multiple Regression

- 4.1 Introduction 41
 - Exercise set A 44
- 4.2 Standard errors 45
 - Things we don't need 49
 - Exercise set B 49
- 4.3 Explained variance in multiple regression 51
 - Association or causation? 53
 - Exercise set C 53
- 4.4 What happens to OLS if the assumptions break down? 53
- 4.5 Discussion questions 53
- 4.6 End notes 59

5 Multiple Regression: Special Topics

- 5.1 Introduction 61
- 5.2 OLS is BLUE 61
 - Exercise set A 63
- 5.3 Generalized least squares 63
 - Exercise set B 65
- 5.4 Examples on GLS 65
 - Exercise set C 66
- 5.5 What happens to GLS if the assumptions break down? 68
- 5.6 Normal theory 68
 - Statistical significance 70
 - Exercise set D 71
- 5.7 The F -test 72
 - "The" F -test in applied work 73
 - Exercise set E 74
- 5.8 Data snooping 74
 - Exercise set F 76
- 5.9 Discussion questions 76
- 5.10 End notes 78

6 Path Models

- 6.1 Stratification 81
 - Exercise set A 86
- 6.2 Hooke's law revisited 87
 - Exercise set B 88
- 6.3 Political repression during the McCarthy era 88
 - Exercise set C 90

6.4	Inferring causation by regression	91
	Exercise set D	93
6.5	Response schedules for path diagrams	94
	Selection vs intervention	101
	Structural equations and stable parameters	101
	Ambiguity in notation	102
	Exercise set E	102
6.6	Dummy variables	103
	Types of variables	104
6.7	Discussion questions	105
6.8	End notes	112
7 Maximum Likelihood		
7.1	Introduction	115
	Exercise set A	119
7.2	Probit models	121
	Why not regression?	123
	The latent-variable formulation	123
	Exercise set B	124
	Identification vs estimation	125
	What if the U_i are $N(\mu, \sigma^2)$?	126
	Exercise set C	127
7.3	Logit models	128
	Exercise set D	128
7.4	The effect of Catholic schools	130
	Latent variables	132
	Response schedules	133
	The second equation	134
	Mechanics: bivariate probit	136
	Why a model rather than a cross-tab?	138
	Interactions	138
	More on table 3 in Evans and Schwab	139
	More on the second equation	139
	Exercise set E	140
7.5	Discussion questions	141
7.6	End notes	150
8 The Bootstrap		
8.1	Introduction	155
	Exercise set A	166

8.2	Bootstrapping a model for energy demand	167
	Exercise set B	173
8.3	End notes	174
9	Simultaneous Equations	
9.1	Introduction	176
	Exercise set A	181
9.2	Instrumental variables	181
	Exercise set B	184
9.3	Estimating the butter model	184
	Exercise set C	185
9.4	What are the two stages?	186
	Invariance assumptions	187
9.5	A social-science example: education and fertility	187
	More on Rindfuss et al	191
9.6	Covariates	192
9.7	Linear probability models	193
	The assumptions	194
	The questions	195
	Exercise set D	196
9.8	More on IVLS	197
	Some technical issues	197
	Exercise set E	198
	Simulations to illustrate IVLS	199
9.9	Discussion questions	200
9.10	End notes	207
10	Issues in Statistical Modeling	
10.1	Introduction	209
	The bootstrap	211
	The role of asymptotics	211
	Philosophers' stones	211
	The modelers' response	212
10.2	Critical literature	212
10.3	Response schedules	217
10.4	Evaluating the models in chapters 7–9	217
10.5	Summing up	218
	References	219
	Answers to Exercises	235

The Computer Labs 294

Appendix: Sample MATLAB Code 310

Reprints

Gibson on McCarthy 315

Evans and Schwab on Catholic Schools 343

Rindfuss et al on Education and Fertility 377

Schneider et al on Social Capital 402

Index 431