

Rainer Winkelmann · Stefan Boes

Analysis of Microdata

Second Edition

 Springer

Contents

1	Introduction	1
1.1	What Are Microdata?	1
1.2	Types of Microdata	4
1.2.1	Qualitative Data	4
1.2.2	Quantitative Data	6
1.3	Why Not Linear Regression?	8
1.4	Common Elements of Microdata Models	11
1.5	Examples	12
1.5.1	Determinants of Fertility	12
1.5.2	Secondary School Choice	16
1.5.3	Female Hours of Work and Wages	18
1.6	Overview of the Book	20
2	From Regression to Probability Models	21
2.1	Introduction	21
2.2	Conditional Probability Functions	23
2.2.1	Definition	23
2.2.2	Estimation	25
2.2.3	Interpretation	26
2.3	Probability and Probability Distributions	30
2.3.1	Axioms of Probability	30
2.3.2	Univariate Random Variables	31
2.3.3	Multivariate Random Variables	32
2.3.4	Conditional Probability Models	36
2.4	Further Exercises	41
3	Maximum Likelihood Estimation	47
3.1	Introduction	47
3.2	Likelihood Function	48
3.2.1	Score Function and Hessian Matrix	50
3.2.2	Conditional Models	52

3.2.3	Maximization	52
3.3	Properties of the Maximum Likelihood Estimator	55
3.3.1	Expected Score	56
3.3.2	Consistency	57
3.3.3	Information Matrix	59
3.3.4	Asymptotic Distribution	61
3.3.5	Covariance Matrix	62
3.4	Normal Linear Model	65
3.5	Further Aspects of Maximum Likelihood Estimation	69
3.5.1	Invariance and Delta Method	69
3.5.2	Numerical Optimization	72
3.5.3	Identification	77
3.5.4	Quasi Maximum Likelihood	78
3.6	Testing	79
3.6.1	Introduction	79
3.6.2	Restricted Maximum Likelihood	82
3.6.3	Wald Test	84
3.6.4	Likelihood Ratio Test	86
3.6.5	Score Test	89
3.6.6	Model Selection	90
3.6.7	Goodness-of-Fit	91
3.7	Pros and Cons of Maximum Likelihood	92
3.8	Further Exercises	93
4	Binary Response Models	97
4.1	Introduction	97
4.2	Models for Binary Response Variables	99
4.2.1	General Framework	99
4.2.2	Linear Probability Model	100
4.2.3	Probit Model	102
4.2.4	Logit Model	104
4.2.5	Interpretation of Parameters	106
4.3	Discrete Choice Models	109
4.4	Estimation	113
4.4.1	Maximum Likelihood	113
4.4.2	Perfect Prediction	116
4.4.3	Properties of the Estimator	117
4.4.4	Endogenous Regressors in Binary Response Models	119
4.4.5	Estimation of Marginal Effects	122
4.5	Goodness-of-Fit	125
4.6	Non-Standard Sampling Schemes	130
4.6.1	Stratified Sampling	130
4.6.2	Exogenous Stratification	130
4.6.3	Endogenous Stratification	131
4.7	Flexible Specification of Binary Response Models	132

4.8	Further Exercises	135
5	Multinomial Response Models	141
5.1	Introduction	141
5.2	Multinomial Logit Model	143
5.2.1	Basic Model	143
5.2.2	Estimation	144
5.2.3	Interpretation of Parameters	148
5.3	Conditional Logit Model	154
5.3.1	Introduction	154
5.3.2	General Model of Choice	155
5.3.3	Modeling Conditional Logits	156
5.3.4	Interpretation of Parameters	159
5.3.5	Independence of Irrelevant Alternatives	163
5.4	Generalized Multinomial Response Models	164
5.4.1	Multinomial Probit Model	165
5.4.2	Mixed Logit Models	168
5.4.3	Nested Logit Models	169
5.5	Further Exercises	170
6	Ordered Response Models	175
6.1	Introduction	175
6.2	Standard Ordered Response Models	178
6.2.1	General Framework	178
6.2.2	Ordered Probit Model	180
6.2.3	Ordered Logit Model	181
6.2.4	Estimation	183
6.2.5	Interpretation of Parameters	183
6.2.6	Single Indices and Parallel Regression	190
6.3	Generalized Threshold Models	192
6.3.1	Generalized Ordered Logit and Probit Models	192
6.3.2	Interpretation of Parameters	193
6.4	Sequential Models	198
6.4.1	Modeling Conditional Transitions	198
6.4.2	Generalized Conditional Transition Probabilities	200
6.4.3	Marginal Effects	201
6.4.4	Estimation	202
6.5	Interval Data	204
6.6	Further Exercises	206
7	Limited Dependent Variables	211
7.1	Introduction	211
7.1.1	Corner Solution Outcomes	212
7.1.2	Sample Selection Models	213
7.1.3	Treatment Effect Models	214

7.2	Tobin's Corner Solution Model	215
7.2.1	Introduction	215
7.2.2	Tobit Model	216
7.2.3	Truncated Normal Distribution	218
7.2.4	Inverse Mills Ratio and its Properties	219
7.2.5	Interpretation of the Tobit Model	222
7.2.6	Comparing Tobit and OLS	225
7.2.7	Further Specification Issues	227
7.3	Sample Selection Models	228
7.3.1	Introduction	228
7.3.2	Censored Regression Model	230
7.3.3	Estimation of the Censored Regression Model	232
7.3.4	Truncated Regression Model	234
7.3.5	Incidental Censoring	235
7.3.6	Example: Estimating a Labor Supply Model	242
7.4	Treatment Effect Models	244
7.4.1	Introduction	244
7.4.2	Endogenous Binary Variable	247
7.4.3	Switching Regression Model	249
7.5	Further Exercises	251
8	Event History Models	255
8.1	Introduction	255
8.2	Duration Models	258
8.2.1	Introduction	258
8.2.2	Basic Concepts	258
8.2.3	Discrete Time Duration Models	263
8.2.4	Continuous Time Duration Models	266
8.2.5	Key Element: Hazard Function	269
8.2.6	Duration Dependence	271
8.2.7	Unobserved Heterogeneity	275
8.3	Count Data Models	283
8.3.1	Poisson Regression Model	283
8.3.2	Unobserved Heterogeneity	288
8.3.3	Efficient versus Robust Estimation	293
8.3.4	Censoring and Truncation	293
8.3.5	Hurdle and Zero-Inflated Count Data Models	295
8.4	Further Exercises	298
	References	301
	Solutions to Selected Exercises	311
	Index	339