

Contents

List of tables	xiii
List of figures	xv
Preface to the fourth edition	xvii
Versions of Stata	xix
Notation and typography	xxi
1 Theory and practice	1
1.1 The likelihood-maximization problem	2
1.2 Likelihood theory	4
1.2.1 All results are asymptotic	8
1.2.2 Likelihood-ratio tests and Wald tests	9
1.2.3 The outer product of gradients variance estimator	10
1.2.4 Robust variance estimates	11
1.3 The maximization problem	13
1.3.1 Numerical root finding	13
Newton's method	13
The Newton–Raphson algorithm	15
1.3.2 Quasi-Newton methods	17
The BHHH algorithm	18
The DFP and BFGS algorithms	18
1.3.3 Numerical maximization	19
1.3.4 Numerical derivatives	20
1.3.5 Numerical second derivatives	24
1.4 Monitoring convergence	25

2	Introduction to ml	29
2.1	The probit model	29
2.2	Normal linear regression	32
2.3	Robust standard errors	34
2.4	Weighted estimation	35
2.5	Other features of method-gf0 evaluators	36
2.6	Limitations	36
3	Overview of ml	39
3.1	The terminology of ml	39
3.2	Equations in ml	40
3.3	Likelihood-evaluator methods	48
3.4	Tools for the ml programmer	51
3.5	Common ml options	51
3.5.1	Subsamples	51
3.5.2	Weights	52
3.5.3	OPG estimates of variance	53
3.5.4	Robust estimates of variance	54
3.5.5	Survey data	56
3.5.6	Constraints	57
3.5.7	Choosing among the optimization algorithms	57
3.6	Maximizing your own likelihood functions	61
4	Method lf	63
4.1	The linear-form restrictions	64
4.2	Examples	65
4.2.1	The probit model	65
4.2.2	Normal linear regression	66
4.2.3	The Weibull model	69
4.3	The importance of generating temporary variables as doubles	71
4.4	Problems you can safely ignore	73

4.5	Nonlinear specifications	74
4.6	The advantages of lf in terms of execution speed	75
5	Methods lf0, lf1, and lf2	77
5.1	Comparing these methods	77
5.2	Outline of evaluators of methods lf0, lf1, and lf2	78
5.2.1	The todo argument	79
5.2.2	The b argument	79
	Using mlevel to obtain values from each equation	80
5.2.3	The lnfj argument	82
5.2.4	Arguments for scores	83
5.2.5	The H argument	84
	Using mlmatsum to define H	86
5.2.6	Aside: Stata's scalars	87
5.3	Summary of methods lf0, lf1, and lf2	90
5.3.1	Method lf0	90
5.3.2	Method lf1	92
5.3.3	Method lf2	94
5.4	Examples	96
5.4.1	The probit model	96
5.4.2	Normal linear regression	98
5.4.3	The Weibull model	104
6	Methods d0, d1, and d2	109
6.1	Comparing these methods	109
6.2	Outline of method d0, d1, and d2 evaluators	110
6.2.1	The todo argument	111
6.2.2	The b argument	111
6.2.3	The lnf argument	112
	Using lnf to indicate that the likelihood cannot be calculated	113
	Using mlsum to define lnf	114

6.2.4	The g argument	116
	Using <code>mlvecsum</code> to define g	116
6.2.5	The H argument	118
6.3	Summary of methods <code>d0</code> , <code>d1</code> , and <code>d2</code>	119
6.3.1	Method <code>d0</code>	119
6.3.2	Method <code>d1</code>	122
6.3.3	Method <code>d2</code>	124
6.4	Panel-data likelihoods	126
6.4.1	Calculating $\ln f$	128
6.4.2	Calculating g	132
6.4.3	Calculating H	136
	Using <code>mlmatbysum</code> to help define H	136
6.5	Other models that do not meet the linear-form restrictions	144
7	Debugging likelihood evaluators	151
7.1	<code>ml check</code>	151
7.2	Using the debug methods	153
7.2.1	First derivatives	155
7.2.2	Second derivatives	165
7.3	<code>ml trace</code>	168
8	Setting initial values	171
8.1	<code>ml search</code>	172
8.2	<code>ml plot</code>	175
8.3	<code>ml init</code>	177
9	Interactive maximization	181
9.1	The iteration log	181
9.2	Pressing the Break key	182
9.3	Maximizing difficult likelihood functions	184
10	Final results	187
10.1	Graphing convergence	187
10.2	Redisplaying output	188

11	Mata-based likelihood evaluators	193
11.1	Introductory examples	193
11.1.1	The probit model	193
11.1.2	The Weibull model	196
11.2	Evaluator function prototypes	198
	Method-lf evaluators	199
	lf-family evaluators	199
	d-family evaluators	200
11.3	Utilities	201
	Dependent variables	202
	Obtaining model parameters	202
	Summing individual or group-level log likelihoods	203
	Calculating the gradient vector	203
	Calculating the Hessian	204
11.4	Random-effects linear regression	205
11.4.1	Calculating lnf	206
11.4.2	Calculating g	207
11.4.3	Calculating H	208
11.4.4	Results at last	209
12	Writing do-files to maximize likelihoods	213
12.1	The structure of a do-file	213
12.2	Putting the do-file into production	214
13	Writing ado-files to maximize likelihoods	217
13.1	Writing estimation commands	217
13.2	The standard estimation-command outline	219
13.3	Outline for estimation commands using ml	220
13.4	Using ml in noninteractive mode	221
13.5	Advice	222
13.5.1	Syntax	223
13.5.2	Estimation subsample	225

13.5.3	Parsing with help from mlopts	229
13.5.4	Weights	232
13.5.5	Constant-only model	233
13.5.6	Initial values	237
13.5.7	Saving results in e()	240
13.5.8	Displaying ancillary parameters	240
13.5.9	Exponentiated coefficients	242
13.5.10	Offsetting linear equations	244
13.5.11	Program properties	246
14	Writing ado-files for survey data analysis	249
14.1	Program properties	249
14.2	Writing your own predict command	252
15	Other examples	255
15.1	The logit model	255
15.2	The probit model	257
15.3	Normal linear regression	259
15.4	The Weibull model	262
15.5	The Cox proportional hazards model	265
15.6	The random-effects regression model	268
15.7	The seemingly unrelated regression model	271
A	Syntax of ml	285
B	Likelihood-evaluator checklists	307
B.1	Method lf	307
B.2	Method d0	308
B.3	Method d1	309
B.4	Method d2	311
B.5	Method lf0	314
B.6	Method lf1	315
B.7	Method lf2	317

C	Listing of estimation commands	321
C.1	The logit model	321
C.2	The probit model	323
C.3	The normal model	325
C.4	The Weibull model	327
C.5	The Cox proportional hazards model	330
C.6	The random-effects regression model	332
C.7	The seemingly unrelated regression model	335
	References	343
	Author index	347
	Subject index	349