

Discrete Choice Methods with Simulation

Kenneth E. Train

University of California, Berkeley

and

National Economic Research Associates, Inc.



*To
Daniel McFadden
and
in memory of
Kenneth Train, Sr.*

Discrete Choice Methods with Simulation

This book describes the new generation of discrete choice methods, focusing on the many advances that are made possible by simulation. Researchers use these statistical methods to examine the choices that consumers, households, firms, and other agents make. Each of the major models is covered: logit, generalized extreme value (including nested and cross-nested logits), probit, and mixed logit, plus a variety of specifications that build on these basics. Simulation-assisted estimation procedures are investigated and compared, including maximum simulated likelihood, the method of simulated moments, and the method of simulated scores. Procedures for drawing from densities are described, including variance reduction techniques such as antithetics and Halton draws. Recent advances in Bayesian procedures are explored, including the use of the Metropolis–Hastings algorithm and its variant Gibbs sampling. No other book incorporates all these topics, which have risen in the past 20 years. The procedures are applicable in many fields, including energy, transportation, environmental studies, health, labor, and marketing.

Professor Kenneth E. Train teaches econometrics, regulation, and industrial organization at the University of California, Berkeley. He also serves as Vice President of National Economic Research Associates (NERA), Inc. in San Francisco, California. The author of *Optimal Regulation: The Economic Theory of Natural Monopoly* (1991) and *Qualitative Choice Analysis* (1986), Dr. Train has written more than 50 articles on economic theory and regulation. He chaired the Center for Regulatory Policy at the University of California, Berkeley, from 1993 to 2000 and has testified as an expert witness in regulatory proceedings and court cases. He has received numerous awards for his teaching and research.

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK
40 West 20th Street, New York, NY 10011-4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa
<http://www.cambridge.org>

© Kenneth E. Train 2003

This book is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without
the written permission of Cambridge University Press.

First published 2003

Printed in the United Kingdom at the University Press, Cambridge

Typeface Times Roman 11/13 pt. *System* L^AT_EX 2_ε [TB]

A catalog record for this book is available from the British Library.

Library of Congress Cataloging in Publication Data

Train, Kenneth.

Discrete choice methods with simulation / Kenneth E. Train.

p. cm.

Includes bibliographical references and index.

ISBN 0-521-81696-3 – ISBN 0-521-01715-7 (pb.)

1. Decision making – Simulation methods. 2. Consumers' preferences –
Simulation methods. I. Title.

HD30.23 .T725 2003

003'.56 – dc21

2002071479

ISBN 0 521 81696 3 hardback

ISBN 0 521 01715 7 paperback

Contents

1	Introduction	<i>page</i> 1
1.1	Motivation	1
1.2	Choice Probabilities and Integration	3
1.3	Outline of Book	7
1.4	Topics Not Covered	8
1.5	A Couple of Notes	11
Part I Behavioral Models		
2	Properties of Discrete Choice Models	15
2.1	Overview	15
2.2	The Choice Set	15
2.3	Derivation of Choice Probabilities	18
2.4	Specific Models	21
2.5	Identification of Choice Models	23
2.6	Aggregation	33
2.7	Forecasting	36
2.8	Recalibration of Constants	37
3	Logit	38
3.1	Choice Probabilities	38
3.2	The Scale Parameter	44
3.3	Power and Limitations of Logit	46
3.4	Nonlinear Representative Utility	56
3.5	Consumer Surplus	59
3.6	Derivatives and Elasticities	61
3.7	Estimation	64
3.8	Goodness of Fit and Hypothesis Testing	71
3.9	Case Study: Forecasting for a New Transit System	75
3.10	Derivation of Logit Probabilities	78

vi Contents

4	GEV	80
4.1	Introduction	80
4.2	Nested Logit	81
4.3	Three-Level Nested Logit	90
4.4	Overlapping Nests	93
4.5	Heteroskedastic Logit	96
4.6	The GEV Family	97
5	Probit	101
5.1	Choice Probabilities	101
5.2	Identification	104
5.3	Taste Variation	110
5.4	Substitution Patterns and Failure of IIA	112
5.5	Panel Data	114
5.6	Simulation of the Choice Probabilities	118
6	Mixed Logit	138
6.1	Choice Probabilities	138
6.2	Random Coefficients	141
6.3	Error Components	143
6.4	Substitution Patterns	145
6.5	Approximation to Any Random Utility Model	145
6.6	Simulation	148
6.7	Panel Data	149
6.8	Case Study	151
7	Variations on a Theme	155
7.1	Introduction	155
7.2	Stated-Preference and Revealed-Preference Data	156
7.3	Ranked Data	160
7.4	Ordered Responses	163
7.5	Contingent Valuation	168
7.6	Mixed Models	170
7.7	Dynamic Optimization	173
Part II Estimation		
8	Numerical Maximization	189
8.1	Motivation	189
8.2	Notation	189
8.3	Algorithms	191
8.4	Convergence Criterion	202
8.5	Local versus Global Maximum	203
8.6	Variance of the Estimates	204
8.7	Information Identity	205

Contents		vii
9	Drawing from Densities	208
9.1	Introduction	208
9.2	Random Draws	208
9.3	Variance Reduction	217
10	Simulation-Assisted Estimation	240
10.1	Motivation	240
10.2	Definition of Estimators	241
10.3	The Central Limit Theorem	248
10.4	Properties of Traditional Estimators	250
10.5	Properties of Simulation-Based Estimators	253
10.6	Numerical Solution	260
11	Individual-Level Parameters	262
11.1	Introduction	262
11.2	Derivation of Conditional Distribution	265
11.3	Implications of Estimation of θ	267
11.4	Monte Carlo Illustration	270
11.5	Average Conditional Distribution	272
11.6	Case Study: Choice of Energy Supplier	273
11.7	Discussion	283
12	Bayesian Procedures	285
12.1	Introduction	285
12.2	Overview of Bayesian Concepts	287
12.3	Simulation of the Posterior Mean	294
12.4	Drawing from the Posterior	296
12.5	Posteriors for the Mean and Variance of a Normal Distribution	297
12.6	Hierarchical Bayes for Mixed Logit	302
12.7	Case Study: Choice of Energy Supplier	308
12.8	Bayesian Procedures for Probit Models	316
	<i>Bibliography</i>	319
	<i>Index</i>	331

