

Inhaltsverzeichnis

Vorwort	V
Inhaltsverzeichnis	IX
Abbildungsverzeichnis	XV
Tabellenverzeichnis	XVII
1 Einleitung	1
1.1 Was sind kategoriale Daten?	1
1.1.1 Einige Beispiele	2
1.1.2 Andere Abgrenzungskriterien	12
1.1.3 Tabellen versus Individualdaten	18
1.1.4 Modelle für kategoriale Daten im Überblick	19
1.1.5 Notation	22
1.2 Analyse kategorialer Daten	23
1.2.1 Einfache deskriptive Techniken	23
1.2.2 Spezifikation: Übersetzung der Untersuchungsfragen in statistische Modelle	29
1.2.3 Datenerhebung: Die Unsicherheit möglicher Stichproben ...	33
1.2.4 Schätzung und Modelltests: Überprüfung statistischer Modelle mit Hilfe empirischer Daten	35
1.2.5 Einige Anwendungsprobleme	49
1.3 Weitere Literaturhinweise	53
2 Tabellenanalyse mit gewichteter Regression: Der GSK-Ansatz	55
2.1 Ein einführendes Beispiel	57
2.1.1 Grundlegende Konzepte und Notation	57
2.1.2 Saturierte Modelle	68

2.1.3	Die Spezifikation von Haupteffekten, Interaktionseffekten und konditionalen Haupteffekten mit Design-Matrizen	78
2.1.4	Nicht-saturierte Modelle	84
2.1.5	Test und Anpassung nicht-saturierter Modelle	95
2.2	Komplexere Datenkonstellationen	112
2.2.1	Logistische Regressionsmodelle	113
2.2.2	Polytome unabhängige Variablen	117
2.2.3	Polytome abhängige Variablen	123
2.2.4	Ordinale und metrische Variablen	127
2.2.5	Kleine Stichproben und Nullzellen	127
2.3	Weitere Möglichkeiten	129
2.3.1	Pfadmodelle	129
2.3.2	Komplizierte Funktionen der Anteilswerte	130
2.3.3	Fehlende Werte, schrittweise Variablenauswahl und komplexe Auswahlverfahren	132
2.4	Anwendungsvoraussetzungen	133
2.5	Literatur- und Programmhinweise	135
3	Log-lineare Analyse kategorialer Daten	137
3.1	Einleitung	137
3.2	Grundlegende Konzepte	139
3.3	Das saturierte Häufigkeitsmodell	146
3.4	Nicht-saturierte Häufigkeitsmodelle	159
3.5	Test und Anpassung nicht-saturierter Modelle	169
3.6	Das Aggregierungstheorem	180
3.7	Das Logitmodell	182
3.7.1	Der modifizierte Regressionsansatz: Logitmodelle mit einer dichotomen abhängigen Variable	183
3.7.2	Der modifizierte Regressionsansatz: Logitmodelle mit einer polytomen abhängigen Variable	187
3.7.3	Modifizierte Pfadanalyse	189
3.8	Ordinale kategoriale Variablen	197
3.9	Einige Probleme: Kleiner Stichprobenumfang und polytome Variablen	204
3.10	Literatur- und Programmhinweise	207
4	Latente Klassenanalyse und log-lineare Modelle mit latenten Variablen	209
4.1	Einleitung	209
4.2	Latente Klassenanalyse: Log-lineare Modelle mit einer latenten Variable	211

4.2.1	Das Basismodell	211
4.2.2	Ermittlung der Maximum-Likelihood-Schätzer: Der EM-Algorithmus	218
4.2.3	Restringierte Modelle	223
4.2.4	Identifizierbarkeit und Tests der Modellanpassung	226
4.3	Assoziationen zwischen latenten und externen Variablen: Latente Klassenwerte und „interne“ Variablen	227
4.4	Modelle mit zwei oder mehr latenten Variablen	233
4.4.1	Saturierte Modelle für latente Variablen	233
4.4.2	Nicht-saturierte Modelle und modifizierte Pfadanalyse- modelle für latente Variablen: Der modifizierte LISREL-Ansatz	238
4.5	Lokale Abhängigkeitsmodelle: Direkte Effekte zwischen Indikatoren und korrelierte Meßfehler	240
4.6	Simultane Analyse in verschiedenen Gruppen	242
4.7	Kausale Modelle mit latenten Variablen	251
4.8	Ordinale latente Variablen	256
4.9	Literatur- und Programmhinweise	258
5	Logistische Modelle für Individualdaten	261
5.1	Ausgangspunkt: Lineare Regression	262
5.2	Das binäre Logitmodell	265
5.2.1	Die Interpretation der Regressionskoeffizienten	267
5.2.2	Das Logitmodell mit mehreren erklärenden Variablen	272
5.2.3	Die Spezifikation nominalskalierter unabhängiger Variablen	276
5.2.4	Statistische Absicherungen der Modellschätzung	280
5.2.5	Die Erklärungskraft eines Logitmodells	287
5.2.6	Modellmodifikation	294
5.3	Logitmodelle für polytome abhängige Variablen	299
5.3.1	Das multinomiale Logitmodell	299
5.3.2	Das konditionale Logitmodell	306
5.3.3	Logitmodelle für ordinale abhängige Variablen	315
5.4	Alternative Modelle	320
5.4.1	Weitere Modelle	321
5.4.2	Alternativen zu logistischen Modellen	322
5.5	Literatur- und Programmhinweise	324
6	Eine GSK-Analyse zum Zusammenhang von objektiven Lebensbedingungen und subjektivem Wohlbefinden	327
6.1	Daten und Modell	328
6.2	Warum ist das ein Beispiel für kategoriale Datenanalyse	331

6.3	Auswahl und Operationalisierung der untersuchten Variablen	333
6.4	Determinanten der Beurteilung der wirtschaftlichen Lage ...	335
6.4.1	Ein einfaches Haupteffektmodell	336
6.4.2	Ein Modell mit konditionalen Effekten	340
6.4.3	Ein Modell mit zwei Subgruppen	342
6.5	Einige Erweiterungen	344
6.5.1	Ordinale abhängige Variablen	344
6.5.2	Metrische unabhängige Variablen	352
6.6	Literaturhinweise	356
7	Eine log-lineare Kohortenanalyse der Religiosität	357
7.1	Das Problem	357
7.2	Kohortenanalyse	359
7.3	Das Alter×Periode-Design	364
7.4	Das Alter×Kohorte-Design	368
7.5	Das Periode×Kohorte-Design	371
7.6	Das Drei-Faktor-Design: Alter×Periode×Kohorte	372
7.7	Ein Vergleich additiver und multiplikativer Modelle	376
7.8	Literaturhinweise	378
8	Eine log-lineare Panelanalyse politischer Präferenzen ...	379
8.1	Einleitung	379
8.2	Analysen auf der manifesten Ebene	381
8.2.1	Veränderungen des Zusammenhangs zwischen Partei- und Kandidatenpräferenz	381
8.2.2	Veränderungen der Partei- und der Kandidatenpräferenz im Zeitablauf	387
8.2.3	Modifizierte Pfadmodelle für Veränderungen der Partei- und Kandidatenpräferenz im Zeitablauf	389
8.3	Modelle mit latenten Variablen	394
8.3.1	Meßmodelle	394
8.3.2	Meßfehler und Veränderungen	396
8.4	Literaturhinweise	401
9	Gibt es den rationalen Wähler? Eine Logitanalyse zur Erklärungskraft des Rational-Choice-Ansatzes in der empirischen Wahlforschung	403
9.1	Die Determinanten des Wahlverhaltens nach der Theorie des rationalen Wählers	404
9.2	Operationalisierung der Konzepte	407
9.3	Logitanalysen von Wahlbeteiligung und Nichtwahl	410

9.4	Simultane Analyse von Wahlbeteiligung und Parteienwahl mit konditionalen Logitmodellen	416
9.5	Mehrstufige Logitmodelle zur Analyse von Wahlbeteiligung und Wahlverhalten	424
9.6	Diskussion	427
9.7	Literatur- und Programmhinweise	428
Anhänge		
A1	Elementare Matrixalgebra	431
A1.1	Was ist eine Matrix?	431
A1.2	Was ist Matrixalgebra?	432
A2	Logarithmus und Exponentialfunktion	435
A3	Chi-Quadrat- und Standardnormalverteilung	438
A4	Datenliste	439
A5	Diskette zum Buch	445
Literaturverzeichnis		447

Abbildungsverzeichnis

Abbildung 1.1:	Modelle zur Analyse kategorialer Daten	20
Abbildung 1.2:	Lineare Regression der Wahlbeteiligung auf das Alter	26
Abbildung 1.3a:	Log-Likelihood-Fläche (unterschiedliche Anteile in Ost und West)	43
Abbildung 1.3b:	Log-Likelihood-Fläche (gleiche Anteile in Ost und West)	44
Abbildung 2.1:	Baumdiagramm mit den Parametereinschätzungen des Modells $[1, B, A B]$	84
Abbildung 2.2:	Regression der Wähleranteile (P1) auf das Alter (X2)	91
Abbildung 3.1:	Das Aggregationstheorem	181
Abbildung 3.2:	Pfaddiagramm für die Daten in Tabelle 1.2	189
Abbildung 4.1:	Das Basismodell latenter Klassen mit einer latenten Variablen X und vier manifesten Variablen A, B, C und D	213
Abbildung 4.2:	Log-lineare Effekte des Basismodells latenter Klassen	219
Abbildung 4.3:	Ein Modell latenter Klassen mit einer „externen“ Variable Generation (G)	231
Abbildung 4.4:	Ein Modell latenter Klassen mit zwei latenten Variablen Y und Z und vier manifesten Variablen A, B, C und D	234
Abbildung 4.5:	Log-lineare Effekte des Basismodells latenter Klassen mit zwei latenten Variablen	237
Abbildung 4.6:	Ein Basismodell latenter Klassen mit direkten Effekten zwischen den Indikatoren	241

Abbildung 4.7:	Verschiedene Beziehungen zwischen den latenten Variablen (X), den manifesten Variablen (M) und der Gruppierungsvariablen (S)	245
Abbildung 4.8:	Log-Lineare Effekte des modifizierten Pfadmodells {YZ,S}, {SYZ, YA, YD, ZB, ZC, SA, SB, SC, SD} . .	249
Abbildung 4.9:	Pfadmodell zur Analyse der Parteipräferenz mit einer latenten Variable und zwei Hintergrundmerkmalen	254
Abbildung 5.1:	Regressionskurve des Logitmodells ($\beta_0 = 0, \beta_1 = 1$)	265
Abbildung 5.2:	Logistische Regression der Wahlbeteiligung auf das Alter	266
Abbildung 5.3:	Auswirkungen unterschiedlicher Regressionsgewichte	268
Abbildung 5.4:	Auswirkungen unterschiedlicher Regressionskonstanten	269
Abbildung 5.5:	Regressionskurven der Logitmodelle aus Tabelle 5.2	275
Abbildung 5.6:	Beobachtete und geschätzte Wählerhäufigkeiten der Modelle M_2 und M_3	297
Abbildung 6.1:	Determinanten der subjektiven Wahrnehmung der eigenen wirtschaftliche Lage	330
Abbildung 7.1:	Log-lineare Effekte (β) auf die Religiosität im (A)lter×(P)eriode-, (A)lter×(K)ohorte- und (P)eriode×(K)ohorte-Design	366
Abbildung 7.2:	Log-lineare Effekte (β) auf die Religiosität im (A)lter×(P)eriode×(K)ohorte-Design	375
Abbildung 7.3:	Vergleich multiplikativer (log-linearer) und additiver Effekte der Kohorte auf die Religiosität im Periode×Kohorte-Design	377
Abbildung 8.1:	Modelle mit kreuzverzögerten Panelkorrelationen	390
Abbildung 8.2:	Verschiedene Meßmodelle	395
Abbildung 8.3:	Log-lineares Kovarianzmodell für die Daten aus Tabelle 8.4	398
Abbildung 8.4:	Kausaldiagramm für die Daten aus Tabelle 8.4 . . .	399
Abbildung A2.1:	Graph der Funktionen $\exp(x)$ und $\ln(x)$	435

Tabellenverzeichnis

Tabelle 1.1:	Wertorientierungen in West- und Ostdeutschland 1990	3
Tabelle 1.2:	Sozio-demographische Struktur von Wählern und Nicht-Wählern	5
Tabelle 1.3:	Wahlbeteiligung und Alter der FDP-Anhänger mit Konfession	6
Tabelle 1.4:	Aufgaben eines modernen Wohlfahrtsstaates (Niederlande 1974)	7
Tabelle 1.5:	Anzahl Selbstmorde pro 1 Mio. Einwohner in Frankreich 1889-1891 (Männer über 20 Jahre)	9
Tabelle 1.6:	Beobachtete und (geschätzte) erwartete Häufigkeiten	36
Tabelle 2.1:	Wahlbeteiligung nach Alter	64
Tabelle 2.2:	Wahlbeteiligung nach Alter unter Kontrolle der Konfession	67
Tabelle 2.3:	Wahlbeteiligung nach Alter und Konfession	69
Tabelle 2.4:	Schätzwerte des saturierten Modells	77
Tabelle 2.5:	Test der Modellanpassung und lineare Kontraste ...	100
Tabelle 2.6:	Schätzwerte des Modells [1,B,A B=1,C]	120
Tabelle 2.7:	Schätzwerte des Modells [1,B] mit der Parteipräferenz als abhängiger Variablen	125
Tabelle 3.1:	Alter und Konfession	142
Tabelle 3.2:	Zusammenhänge zwischen drei Variablen	147
Tabelle 3.3:	Schätzwerte für das saturierte Modell	152
Tabelle 3.4:	Modifizierte Pfadanalyse für das Modell in Abbildung 3.2	191
Tabelle 3.5:	Bildungsniveau und Präferenz für eine „linke“ Partei	199
Tabelle 4.1:	Ein Modell mit einer latenten Variable und zwei latenten Klassen für die Daten über notwendige Staatsaufgaben (Niederlande)	217

Tabelle 4.2:	Ein Modell mit zwei latenten Variablen für die Daten über notwendige Staatsaufgaben (Niederlande)	235
Tabelle 4.3:	Aufgaben eines modernen Wohlfahrtsstaates (Deutschland und Schweiz)	243
Tabelle 4.4:	Ein Modell mit zwei latenten Variablen für die Daten über notwendige Staatsaufgaben (Deutschland und Schweiz)	246
Tabelle 4.5:	Religiosität, Einkommen, Parteipräferenz und notwendige Staatsaufgaben (Deutschland)	250
Tabelle 4.6:	Ein Modell mit zwei latenten Variablen für die Daten über notwendige Staatsaufgaben (Deutschland)	251
Tabelle 4.7:	Ein lineares Modell mit einer latenten Variable und drei latenten Klassen für die Daten über notwendige Staatsaufgaben (Niederlande)	257
Tabelle 5.1:	Wahrscheinlichkeiten, Logits und Odds des Logitmodells aus Abbildung 5.2	270
Tabelle 5.2:	Analyse der Wahlbeteiligung mit den unabhängigen Variablen Alter und Konfession	273
Tabelle 5.3:	Analyse der Wahlbeteiligung mit den unabhängigen Variablen Alter und Konfession: Logitmodell mit Interaktionseffekt	276
Tabelle 5.4:	Bildung von Designvariablen für die Parteipräferenz	277
Tabelle 5.5:	Analyse der Wahlbeteiligung mit der unabhängigen Variablen Parteipräferenz	278
Tabelle 5.6:	Die Bedeutung der Koeffizienten bei der Dummy- und der Effektkodierung	279
Tabelle 5.7:	Analyse der Wahlbeteiligung mit den unabhängigen Variablen Alter, Konfession und Parteipräferenz	280
Tabelle 5.8:	Regressionskoeffizienten, Standardschätzfehler, Effekte und 95%-Konfidenzintervalle für Modell M_6	282
Tabelle 5.9:	Varianzen und Kovarianzen der Regressionskoeffizienten von Modell M_6	284
Tabelle 5.10:	Vergleich verschiedener Logitmodelle	290
Tabelle 5.11:	Klassifikationstabelle für das Logitmodell M_6	292
Tabelle 5.12:	Vorhersageerfolgstabelle und Wahrscheinlichkeitsprofile für Modell M_6	293
Tabelle 5.13:	Regressionskoeffizienten, Standardschätzfehler, Effekte und deren 95%-Konfidenzintervalle für Modell M_8	295
Tabelle 5.14:	Analyse der Parteipräferenz durch ein multinomiales Logitmodell mit der unabhängigen Variablen Alter	303

Tabelle 5.15:	Umrechnung von der Referenzkategorie CDU auf andere Referenzkategorien	305
Tabelle 5.16:	Der Einfluß von Alter, Konfession und Bildung auf die Parteipräferenz	307
Tabelle 5.17:	Der Einfluß der Parteipräferenz auf die Wahlabsicht (Spezifikation als generische Variable)	311
Tabelle 5.18:	Der Einfluß der Parteipräferenz auf die Wahlabsicht (Spezifikation über alternativenspezifische Variablen)	313
Tabelle 5.19:	Der Einfluß von Konfession und Bildung auf die Parteipräferenz	314
Tabelle 5.20:	Der Einfluß von Alter auf den Bildungsabschluß im multinomialen Logitmodell	316
Tabelle 5.21:	Der Einfluß von Alter auf den Bildungsabschluß im kumulierten Logitmodell	318
Tabelle 5.22:	Programme zur Schätzung von Logitmodellen	325
Tabelle 6.1:	Wirtschaftliche Lage nach Einkommen, Vergleich mit Vorjahr und Befragungsort	336
Tabelle 6.2:	Durchschnittliche Beurteilung nach Einkommen, Vergleich mit Vorjahr und Befragungsort	345
Tabelle 6.3:	Kumulative Logits nach Einkommen, Vergleich mit Vorjahr und Befragungsort	349
Tabelle 6.4:	WLS-Schätzung für Modell M_6	351
Tabelle 6.5:	Wirtschaftliche Lage nach Einkommensquartil und Befragungsort	353
Tabelle 7.1:	Geschätzte Anteile (in Prozent) der Frauen ohne Konfession (N) nach Alter (A), Periode (P) und Kohorte (K) für die Niederlande	360
Tabelle 7.2:	Mitgliedschaft in einer Konfession nach Periode, Alter und Kohorte	362
Tabelle 7.3:	Teststatistiken des Alter×Periode-Designs	365
Tabelle 7.4:	Teststatistiken des Alter×Kohorte-Designs	369
Tabelle 7.5:	Teststatistiken des Periode×Kohorte-Designs	371
Tabelle 8.1:	Partei- und Kandidatenpräferenz in Deutschland zum Zeitpunkt T_1 (Nov./Dez. 1989) und T_2 (Mai/Juni 1990)	380
Tabelle 8.2:	Der Zusammenhang zwischen Partei- und Kandidatenpräferenz	382
Tabelle 8.3:	Veränderungen der Partei- und der Kandidatenpräferenz im Zeitablauf	388
Tabelle 8.4:	Auswirkungen der Fernseh-Dokumentationsserie auf die Sympathien für neofaschistische Parteien	396

Tabelle 8.5:	Veränderungen der Sympathien für neofaschistische Parteien unter den regelmäßigen Zuschauern der Dokumentationsserie	397
Tabelle 8.6:	Ein Modell latenter Klassen für die Wirkungen der Fernsehserie	400
Tabelle 9.1:	Binäre Logitmodelle zur Prognose der Wahlbeteiligung	413
Tabelle 9.2:	Ergebnisse des binären und des ordinalen Logitmodells	415
Tabelle 9.3:	Ergebnisse der konditionalen und mehrstufigen Logitanalysen	421
Tabelle A2.1:	$\exp(x)$ und $\ln(x)$ für einige ausgewählte Werte	436