# Modelle für kategoriale Daten mit ordinalem Skalenniveau

## Parametrische und nonparametrische Ansätze

von

Gerhard Tutz

Mit 4 Abbildungen und 12 Tabellen

Technische Hochschule Darmstadt
Volkswirtschaftliche Bibliothek

Vandenhoeck & Ruprecht in Göttingen

# Inhaltsverzeichnis


1. Messtheoretische Grundlagen und statistische Invarianzkonzepte . 1
   - 1.1 Fundamentale Messung ..... 1
     - 1.1.1 Skalen und Repräsentation ..... 1
     - 1.1.2 Eindeutigkeit und Skalenniveau ..... 2
     - 1.1.3 Bedeutsamkeit von Skalen ..... 5
   - 1.2 Statistische Invarianzkonzepte und Bedeutsamkeit ..... 9
     - 1.2.1 Invariante Statistiken ..... 9
     - 1.2.2 Invariante Familien von Verteilungen ..... 11
     - 1.2.3 Invarianzeigenschaften des Unabhängigkeitsmodells ..... 12
   - 1.3 Ordinale Meßstrukturen ..... 13
     - 1.3.1 Schwache und einfache Ordnung ..... 13
     - 1.3.2 Strikte Ordnungen ..... 15

2. Zusammenhangsanalyse für kategorial-ordinale Merkmale ..... 17
   - 2.1 Grundlegende Modelle und Invarianzeigenschaften ..... 17
     - 2.1.1 Zweidimensionale loglineare Modelle ..... 17
     - 2.1.2 Parameterisierung des Interaktionsterms im Assoziationsmodell 19
     - 2.1.3 Invarianzeigenschaften des Assoziationsmodells ..... 21
     - 2.1.4 A priori-Wahl des Scores ..... 25
     - 2.1.5 Invarianzeigenschaften bei Wertzuweisenungen ..... 26
   - 2.2 Familie der Assoziationsmodelle ..... 33
     - 2.2.1 Das allgemeine Assoziationsmodell ..... 33
     - 2.2.2 Hierarchische Verschachtelung der Modelle ..... 40
     - 2.2.3 Rang-Assoziations-Modelle ..... 47
   - 2.3 Korrelationsmodelle ..... 48
   - 2.4 Dreidimensionale Modelle ..... 50
     - 2.4.1 Bedingte Assoziationsmodelle ..... 50
     - 2.4.2 Partielle Assoziationsmodelle ..... 51
   - 2.5 Schätz- und Testtheorie für Multinomialverteilungsmodelle ..... 54
     - 2.5.1 Modelldarstellungen ..... 54
     - 2.5.2 Minimum-Distanz-Schätzer bei vollständig parameterisierten Modellen ..... 56
     - 2.5.3 Schätzung für Restriktionsmodelle ..... 58
     - 2.5.4 Gewichtete Kleinst-Quadrat-Schätzung ..... 58
     - 2.5.5 Asymptotisches Verhalten der Schätzer ..... 60
     - 2.5.6 Anpassungstests ..... 63

2.6 Schätzen und Testen für spezielle Modelle ..... 65
2.6.1 Assoziationsmodelle ..... 65
2.6.2 Das Korrelationsmodell ..... 67

3. Regressionsmodelle mir ordinaler Zielgrösse ..... 68
3.1 Kategorial-ordinale Struktur bei Regressionsmodellen ..... 69
3.1.1 Elementare Invarianzpostulate ..... 69
3.1.2 Allgemeine Invarianzpostulate ..... 71
3.1.3 Das Prinzip der maximalen Random Utility ..... 74
3.2 Das kumulative Modell ..... 76
3.2.1 Latente Variablen und Modellvorstellungen ..... 76
3.2.2 Familie der kumulativen Modelle und Zulässigkeit der Parameterisierungen ..... 78
3.2.3 Äquivalenz kumulativer Modelle ..... 81
3.2.4 Invarianzeigenschaften kumulativer Modelle ..... 84
3.2.5 Spezielle Modelle ..... 85
3.3 Das sequentielle Modell ..... 86
3.3.1 Der sequentielle Reaktionsmechanismus der latenten Variablen . 86
3.3.2 Invarianzeigenschaften sequentieller Modelle ..... 89
3.3.3 Beziehungen zwischen sequentiellen und kumulativen Modellen 90
3.3.4 Das inverse sequentielle Modell ..... 93
3.4 Zwei-Schritt-Modelle ..... 96
3.5 Alternative Ansätze ..... 97
3.5.1 Nachbarschaftsmodelle ..... 97
3.5.2 Verallgemeinerte Extremwertmodelle ..... 100
3.6 Aggregierbarkeit der abhängigen Variable ..... 101
3.7 Schätzung mit GLAMOUR ..... 103

4. Multivariate ordinale Regressionsmodelle ..... 105
4.1 Invarianzeigenschaften multivariater Modelle ..... 106
4.2 Modelle mit latenten Variablen ..... 107
4.2.1 Der kumulative Ansatz ..... 107
4.2.2 Strukturgleichungs-Ansätze ..... 108
4.3 Bedingte Parameterisierungen ..... 111
4.3.1 Sukzessive Modellierung der abhängigen Variablen ..... 111
4.3.2 Hypothesen und Modellhierarchie ..... 112
4.3.3 Beziehungen zu den Zwei-Schritt-Modellen ..... 113
4.3.4 Variablen unterschiedlichen Skalenniveaus ..... 116
4.4 Die Parameterisierung ordinaler Assoziationsmasse ..... 116
4.4.1 Globales Kreuzproduktverhältnis ..... 116
4.4.2 Lokales Kreuzproduktverhältnis ..... 118

5. Schätzen und Testen für kategoriale Regressionsmodelle ..... 120
5.1 Verallgemeinerte lineare Modelle ..... 120
5.1.1 Modelldarstellung ..... 120
5.1.2 Maximum-Likelihood-Schätzung ..... 122
5.1.3 Kategoriale Zeitreihen und Dynamisierung des Modells ..... 129
5.1.4 Verallgemeinerte lineare Modelle mit zusammengesetzter Link- Funktion ..... 130
5.1.5 Spezielle Modelle ..... 133
5.2 Methoden für faktorielle probabilistische Reaktionsmodelle mit Meßwiederholungen ..... 136
5.2.1 Charakterisierung der Modellklasse ..... 136
5.2.2 Minimum-Distanz-Schätzer ..... 137
5.2.3 Asymptotik ..... 139

6. Dichteschätzung für kategoriale Variablen ..... 140
6.1 Diskrete Kernschätzung ..... 140
6.2 Invarianzeigenschaften ..... 145
6.3 Ordinale Kerne ..... 147
6.3.1 Ordinaler Aitchison & Aitken-Kern ..... 148
6.3.2 Habbemas Kern ..... 148
6.3.3 Kerne mit unendlichem Träger ..... 149
6.3.4 Gemeinsame Eigenschaften von Kernen ..... 151
6.4 Variable Kerne ..... 151
6.5 Wahl des Glättungsparameters ..... 154
6.5.1 Methoden der allgemeinen Schadensfunktion ..... 154
6.5.2 Asymptotische Aussagen ..... 156

7. Nonparameterische Schätzung abhängiger kategorialer Merkmale 159
7.1 Der diskriminanzanalytische Ansatz ..... 160
7.1.1 Optimale Entscheidungsregeln ..... 160
7.1.2 Geschätzte Entscheidungsregeln ..... 162
7.1.3 Schätzung des Schadens ..... 164
7.2 Diskriminanzfunktion und indirekte Kerne ..... 167
7.2.1 Kreuzvalidierte Wahl der Glättungsparameter für allgemeine Schadensfunktionen ..... 167
7.2.2 Separate Wahl in den Klassen und klassenspezifische Schadensfunktionen ..... 171
7.2.3 Die Familie der zuordnungsspezifischen Schadensfunktionen ... 172
7.2.4 Asymptotisches Verhalten der Schätzer ..... 175

7.3 Direkte Kernschätzmethode ..... 183
7.3.1 Direkte Schätzung der a-posteriori-Wahrscheinlichkeit ..... 184
7.3.2 Direkte Kernschätzung mit distanzabhängiger Glättung ..... 185
7.3.3 Wahl der Glättungsparameter durch direkte Kreuzvalidierung . 190
7.3.4 Wahl der Glättungsparameter durch indirekte Kreuzvalidierung 193
7.3.5 Asymptotische Eigenschaften direkter Kernschätzer ..... 195

8. Nonparametrische geglättete Regression mit kategorialer und ordinaler Zielgrösse ..... 198
8.1 Methoden für kategoriale Zielgröße ..... 198
8.2 Die Methode der direkten Kerne ..... 199
8.3 Modifikation des Schätzverfahrens ..... 203

Appendix 1 ..... 206
Appendix 2 ..... 206
Appendix 3 ..... 207
Notation ..... 208
Literaturverzeichnis ..... 209
Programmpakete für verallgemeinerte lineare Modelle ..... 220
Zusammenfassung ..... 221
Summary ..... 223
Résumé ..... 224
Personenverzeichnis ..... 225
Stichwortverzeichnis ..... 227