MIKROÖKONOMETRIE – DISKRETE ENTSCHEIDUNGSMODELLE

Univ.-Prof. Dr. Joachim Merz

Skriptum zur Vorlesung

Dreizehnte Auflage 2016

Leuphana Universität Lüneburg

Fakultät Wirtschaftswissenschaften

Forschungsinstitut Freie Berufe

Professur 'Statistik und Freie Berufe'

www.leuphana.de/ffb

Copyright © 2016

Dreizehnte Auflage 2016

Univ.-Prof. Dr. Joachim Merz, LEUPHANA Universität Lüneburg, Fakultät Wirtschaftswissenschaften, Forschungsinstitut Freie Berufe (FFB), Professur 'Statistik und Freie Berufe', Campus, Geb. 4, Scharnhorststr. 1, 21335 Lüneburg, Tel.: 04131/677-2051, Fax: 04131/677-2059, E-Mail: merz@uni.leuphana.de, www.leuphana.de/ffb.

Generelle Informationen zur empirischen Wirtschaftsforschung und zu ‚Mikroökonometrie – Diskrete Entscheidungsmodelle' finden Sie auf unserer FFB-Homepage

www.leuphana.de/ffb

Mikroökonometrie – Diskrete Entscheidungsmodelle

Skriptum zur Vorlesung

Univ.-Prof. Dr. Joachim Merz

THEMENBEREICHE

I EINLEITUNG

II BINÄRE ENTSCHEIDUNGSMODELLE

III MULTIPLE ENTSCHEIDUNGSMODELLE

IV MODELLE MIT BEGRENZT ABHÄNGIGEN VARIABLEN

V PANELMODELLE

PROGRAMMPAKETE

LITERATUR

TESTTABELLEN

MIKROÖKONOMETRIE – DISKRETE ENTSCHEIDUNGSMODELLE

I EINLEITUNG 1

1 Mikroökonometrie – Ökonometrie individuellen Verhaltens 1

2 Modelle mit diskreten abhängigen Variablen: 'Discrete Choice' und 'Limited Dependent Variables (LDV)' 6

3 Anwendungsfälle und Beispiele 7

II BINÄRE ENTSCHEIDUNGSMODELLE 9

1 Struktur binärer Entscheidungsmodelle 9

1.1 Binärer Regressionsansatz 9
1.2 Latentes Regressions-Index Funktionsmodell 9
1.3 RUM: Random Utility Model 11

2 Alternative Modelle 14

2.1 Das lineare Wahrscheinlichkeitsmodell: LP 14
2.2 Nichtlineare Transformation I: LOGIT 18
2.3 Nichtlineare Transformation II: PROBIT 22
2.4 Weitere binäre Entscheidungsmodelle: Arctan- gestutzt exponentielle Modelle und weitere Verteilungen 25
2.5 Zusammenfassung und Vergleich LP, LOGIT und PROBIT 28
2.6 Generalisierte Lineare Modelle (GLM) und binäre Entscheidungsmodelle 29
2.7 Umrechnungder geschätzten Koeffizienten zwischen den alternativen Ansätzen LP, LOGIT und PROBIT 31
2.8 Interpretation der Modellergebnisse – Beispiel Kreditvergabe 31

3 Schätzen von binären Entscheidungsmodellen 38

3.1 Schätzen mit gruppierten Daten 38
3.2 Schätzen mit Individualdaten 42
3.3 Exkurs: Newton-Raphson-Methode zur Lösung nichtlinearer Optimierungsprobleme 48

4 Hypothesentests, Modellgüte und Interpretation 54

4.1 Hypothesentests 54
4.2 Gütemaße und Spezifikationstests 55
4.3 Interpretation, Ableitungen und Elastizitäten 57
4.4 Stata, LIMDEP und GAUSS: Gütemaße, Spezifikationstests 58

5 Nutzentheoretische Grundlagen und Binäres Entscheidungsmodell 65

5.1 Mikroökonomische Entscheidungstheorie 65
5.2 Mikroökonomisches Beispiel zur Dualität 66
5.3 Mikroökonomische Nutzenmaximierung und stochastische Nutzenfunktion (RUM: Random Utility Maximization) 69

6 Beispiele 73

6.1 LOGIT Schätzung mit stata – Video example 73
6.2 Training – LOGIT und PROBIT Schätzung mit stata 73
6.3 Leaving the U.S. Navy – Gruppierte und individuelle Schätzungen LP und LOGIT 74
6.4 Nachfrage nach dauerhaften Konsumgütern – Gruppierte Schätzungen LOGIT 77
6.5 Travelling by Auto or Transit – Individualschätzung LOGIT 80

6.6 Arbeitsangebot von Frauen – PROBIT 82

III MULTIPLE ENTSCHEIDUNGSMODELLE 86

1 Struktur multipler Entscheidungsmodelle 86

2 Nutzentheoretische Fundierung: RUM – Random Utility Maximization 86

3 Das MULTINOMIALE LOGIT (MNL-)Modell 88

4 McFadden's CONDITIONAL LOGIT 90

5 Schätzung des MULTINOMIAL LOGIT und des CONDITIONAL LOGIT Modells 91

6 Unabhängigkeit von irrelevanten Alternativen: IIA-Annahme 93

7 LIMDEP: Marginale Effekte und vorhergesagte Wahrscheinlichkeiten 94

8 Beispiele multipler Entscheidungsmodelle 97

8.1 MULTINOMIALES LOGIT / LIMDEP: Berufliche Stellung 97
8.2 MULTINOMIALES LOGIT / Stata: Berufliche Stellung 103
8.3 MULTINOMIALES LOGIT: Sind Selbständige zeit- und einkommensarm? 107
8.4 CONDITIONAL LOGIT: Travel mode choice 109

9 NESTED LOGIT und HETEROSKEDASTISCHES LOGIT Modell 112

9.1 NESTED LOGIT Modell 112
9.2 HETEROSKEDASTISCHES LOGIT (HEV) Modell 114
9.3 Beispiel NESTED und HETEROSKEDASTISCHES LOGIT Modell 114

10 Modelle für geordnete Kategorien: 'Ordered response models' 121

10.1 LOGIT- und PROBIT-Modelle für geordnete Kategorien 121
10.2 GENERALIZED ORDERED PROBIT (GOP) Modell 124
10.3 Beispiele 'Ordered response models' 125

11 Das POISSON-Modell für Zähldaten 134

11.1 Modellspezifikation und Schätzung 134
11.2 Beispiele: POISSON-Modell 135

12 Simultane diskrete Entscheidungsmodelle 138

12.1 Modellspezifikationen 138
12.2 BIVARIATES und MULTIVARIATES PROBIT 138
12.3 Beispiel: BIVARIATES PROBIT - Voting Behaviour 140
12.4 Beispiel: MULTIVARIATES PROBIT - The Timing of Daily Demand for Goods and Services 141

IV MODELLE MIT BEGRENZT ABHÄNGIGEN VARIABLEN 142

1 Modellstruktur Limited dependent variable (LDV)-Modelle: truncation, censoring 142

2 Das TOBIT-Modell 144

2.1 Modellspezifikation 144
2.2 Interpretation des TOBIT-Modells – Zur Frage prinzipiell nicht negativer Beobachtungen (z.B. Ausgaben, Zeitverwendung etc.): 146
2.3 ML-Schätzung 149
2.4 Gütemaße und Spezifikationstests 154
2.5 Beispiele 154

3 Selektion und Regression - Das zweistufige Heckman Modell 160

3.1 Modellspezifikation, selektionskorrigierte Schätzung 160
3.2 Beispiele 161

4 Selektion und Regression – Das generalisierte zweistufige Modell von Lee 169

4.1 Struktur und Schätzung des Ansatzes 169
4.2 Beispiele Lee Ansatz 169

5 Semiparametrische selektionskorrigierte Schätzung 178

5.1 Modellspezifikation, semiparametrische Schätzung 178
5.2 Beispiel: A Comparison of Binary Choice Estimators 179
6 Fractional Regression Models 180

V PANELMODELLE 181

1 Modellstruktur der ‘discrete choice’-Panelmodelle 181
2 Fixed effects ‘discrete choice’ LOGIT und PROBIT Panelmodelle 181
3 Random effects ‘discrete choice’ LOGIT und PROBIT Panelmodelle 182
4 Fixed effects vs. Random effects LOGIT und PROBIT Panelmodelle 182

PROGRAMMPAKETE 184

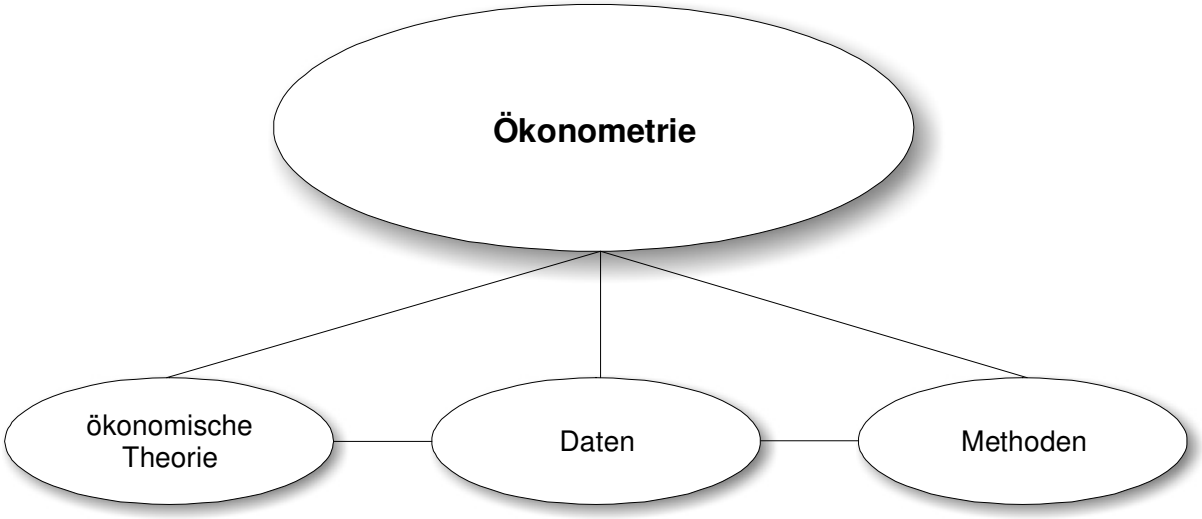
LITERATUR 185

VERTEILUNGSTABELLEN 190

VERZEICHNIS WICHTIGER SYMBOLE 202

I EINLEITUNG

1 Mikroökonometrie – Ökonometrie individuellen Verhaltens



Mikroökonometrie:

- Mikroökonomische Theorie:
  z.B. Theorie des Haushalts, Theorie der Unternehmung, Individuelle Entscheidungen
- Mikrodaten:
  Umfragen, Statistische Erhebungen: Mikrozensus, Einkommens- und Verbrauchsstichprobe EVS, Sozio-ökonomisches Panel etc.
- 'Mikro'-Makro-Methoden:
  Zwar vieles gleich aber seit den 70er Jahren liegen besondere Schwerpunkte auf: 'Discrete Choice'; 80er/90er-Jahre: Panelanalysen
- Mikrosimulation:
  Für die Wirkungsanalyse wirtschafts- und sozialpolitischer und betrieblicher alternativer Politiken

Econometrics

In the first issue of Econometrica, the Econometric Society stated that

- Its main object shall be to promote studies that aim at a unification of the theoretical-quantitative and the empirical-quantitative approach to economic problems and that are penetrated by constructive and rigorous thinking similar to that which has come to dominate the natural sciences. But there are several aspects of the quantitative approach to economics, and no single one of these aspects, taken by itself, should be confounded with econometrics. Thus, econometrics is by no means the same as economic statistics. Nor is it identical with what we call general economic theory, although a considerable portion of this theory has a definitely quantitative character. Nor should economics be taken as synonymous [sic] with the application of mathematics to economics. Experience has shown

that each of these three view-points [sic], that of statistics, economic theory, and mathematics, is a necessary, but not by itself a sufficient, condition for a real understanding of the quantitative relations in modern economic life. It is the unification of all three that is powerful. And it is this unification that constitutes econometrics. (Frisch 1933)

- …econometrics may be defined as the quantitative analysis of actual economic phenomena based on the concurrent development of the theory and observation, related by appropriate methods of inference (Samuelson/Koopmans/Stone 1954, S. 141).
- Econometrics may be defined as the social science in which the tools of economic theory, mathematics, and statistical inference are applied to the analysis of economic phenomena (Goldberger 1964, S. 1).
- Econometrics consists of an application of statistical methods to economic date (Maddala 1977, S. 3).
- Models play a major role in all econometric studies, whether theoretical or applied. Indeed, defining econometrics as the branch of economics concerned with the empirical estimation of economic relationships, models, together with data, represent the basic ingredients of any economic study. Typically, the theory of phenomena under investigation is developed into a model which is further refined into an econometric model (Intrilligator 1983, S. 182).

Mikroökonometrie

Ökonometrie individuellen Verhaltens (Mikroeinheiten; Personen, Haushalte, Familien, Betriebe, Organisationen...)

Die besondere Bedeutung der Mikroökonometrie innerhalb der empirischen Wirtschaftsforschung wird unterstrichen durch die Verleihung des Nobelpreises 2000 an Daniel McFadden ('discrete choice') und James J. Heckman (‚selected samples'):

Press Release - The 2000 Sveriges Riksbank (Bank of Sweden) Prize in Economic Sciences in Memory of Alfred Nobel

KUNGL. VETENSKAPSAKADEMIEN
THE ROYAL SWEDISH ACADEMY OF SCIENCES

October 11, 2000

The Royal Swedish Academy of Sciences has decided that the Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel, 2000, will be shared between

James J. Heckman
University of Chicago, USA, and

Daniel L. McFadden
University of California, Berkeley, USA.

In the field of microeconometrics, each of the laureates has developed theory and methods that are widely used in the statistical analysis of individual and household behavior, within economics as well as other social sciences.

Citation of the Academy:
"to James Heckman for his development of theory and methods for analyzing selective samples and to Daniel McFadden for his development of theory and methods for analyzing discrete choice. "

Microeconometrics - on the boundary between economics and statistics - is a methodology for studying micro data, i.e., economic information about large groups of individuals, households, or firms. Greater availability of micro data and increasingly powerful computers have enabled empirical studies of many new issues. For example, what determines whether an individual decides to work and, if so, how many hours? How do economic incentives affect choices of education, occupation, and place of residence? What are the effects of different educational programs on income and employment? James Heckman and Daniel McFadden have resolved fundamental problems that arise in the statistical analysis of micro data. The methods they have developed have solid foundations in economic theory, but have evolved in close interplay with applied research on important social problems. They are now standard tools, not only among economists but also among other social scientists.

Available micro data often entail **selective samples**. Data on wages, for instance, cannot be sampled randomly if only individuals with certain characteristics - unobservable to the researcher - choose to work or engage in education. If such selection is not taken into account, statistical estimation of economic relationships yields biased results. Heckman has developed statistical methods of handling selective samples in an appropriate way. He has also proposed tools for solving closely related problems with individual differences unobserved by the researcher; such problems are common, e.g. when evaluating social programs or estimating how the duration of unemployment affects chances of getting a job. Heckman is also a leader of applied research in these areas.

Micro data often reflect **discrete choice**. For instance, data regarding individuals' occupation or place of residence reflect choices they have made among a limited number of alternatives. Prior to McFadden's contributions, empirical studies of such choices lacked a foundation in economic theory. Evolving from a new theory of discrete choice, the statistical methods developed by McFadden have transformed empirical research. His methods are readily applicable. For example, they prevail in models of transports and are used to evaluate changes in communication systems. Examples of McFadden's extensive applications of his own methods include the design of the San Francisco BART system, as well as investments in phone service and housing for the elderly.

***

James J. Heckman (US citizen), 56, was born in Chicago, IL in 1944. Since 1995 he is the Henry Schultz Distinguished Service Professor of Economics at the University of Chicago.

Daniel L. McFadden (US citizen), 63, was born in Raleigh, NC in 1937. Since 1990 he holds the E. Morris Cox Chair in Economics at the University of California, Berkeley.

The Prize amount, SEK 9 million, will be shared equally between the Laureates.

Nobelpreis 2000 an James J. Heckman und Daniel McFadden (Mikroökonometrie: discrete choice, selected samples)

Nobelpreis für Wirtschaft geht an „stille Forscher"

US-Wissenschaftler Heckman und McFadden beschäftigen sich mit Motiven ökonomischer Entscheidungen

STOCKHOLM/FRANKFURT A.M. (afp/ap/dpa/msv). Der mit zwei Millionen Mark dotierte Nobelpreis für Wirtschaft geht in diesem Jahr an die Amerikaner James Heckman und Daniel McFadden (beide Bilder: dpa). Die Königlich-schwedische Akademie der Wissenschaften zeichnet die Ökonomen für ihre Forschungen über die Beweggründe wirtschaftlicher Entscheidungen aus. Sie hätten „im Felde der Mikroökonometrie Theorien und Methoden entwickelt, die heute weit reichende Verwendung in der empirischen Analyse des Verhaltens von Individuen und Haushalten finden". Dies gelte außer für Wirtschafts- auch für andere Sozialwissenschaften, erklärte die Akademie. Die Auszeichnung für Wirtschaft wird zusammen mit den anderen Nobelpreisen am 10. Dezember überreicht.

Heckman und McFadden hätten fundamentale Probleme gelöst, die sich bei der Analyse so genannter Mikrodaten ergäben, hieß es weiter. Im Blickpunkt stehen Fragen wie: Was bestimmt die Entscheidung eines Menschen zu arbeiten? Wie beeinflussen wirtschaftliche Anreize Entscheidungen über Ausbildung, Beruf und Wohnort? Wie wirkt sich die Ausbildung auf Einkommen und Beschäftigung aus? Mit den Forschungen können etwa sozialpolitische Hilfsprogramme bewertet werden oder die Chancen von so genannten Langzeitarbeitslosen auf eine Rückkehr ins Berufsleben.

Der 56-jährige Heckman lehrt an der Universität von Chicago (US-Bundesstaat Illinois). Er forschte zuletzt unter anderem über Sozialsysteme, Arbeitsmärkte und alternative Einkommensverteilung. Mit Hilfe seiner Modelle könnten etwa Unterschiede zwischen den Gehältern von Frauen und Männern unter Berücksichtigung ihrer Bildung und des Alters studiert werden, sagte das Mitglied der Akademie, Karl-Gustav Joereskog. Heckmans Studien seien ein gutes Beispiel für die Integration von Theorie und Praxis, insbesondere wenn es um die Planung der Arbeitsmarktpolitik gehe, betonte Akademie-Mitglied Jorgen Weibull.

Ohne Heckmann wären Verfahren, wie sie die Bundesanstalt für Arbeit nutze, undenkbar, meint Rainer Winkelmann vom Forschungsinstitut zur Zukunft der Arbeit.

James Heckman

Als Ökonom der Chicago-Schule schwöre Heckmann auf den freien Markt und betone den negativen Einfluss von staatlichen Eingriffen in den Arbeitsmarkt.

Der 63-jährige McFadden von der kalifornischen Elite-Universität Berkeley erhält den Preis für „die Entwicklung von Theorien und Methoden zur Analyse diskreter Wahlentscheidungen". Seine Methoden wurden etwa für den Aufbau des regionalen Bahnsystems Bart in San Francisco genutzt, bei Telefonangeboten oder der Gestaltung von Seniorenwohnungen.

Während die Nobelpreise für Literatur, Medizin, Physik, Chemie und Frieden seit 1901 verliehen werden, gibt es die Auszeichnung für Wirtschaftswissenschaftler erst seit 1969. Von den bisher 44 Gewinnern kommen 28 aus den Vereinigten Staaten. Eine Frau wurde bisher noch nicht gekürt.

Der Bremer Ökonom Rudolf Hickel lobt das diesjährige Votum als eine „gute Entscheidung", die der Mikroökonomie Impulse geben werde. Die Analysen von Heckmann haben Hickel zufolge bei der Bewertung der Arbeitsmarktpolitik in Deutschland bisher kaum Berücksichtigung gefunden. Das könnte sich ändern, prognostiziert Hickel.

McFaddens Arbeiten lieferten Hinweise auf die Motive, den öffentlichen Personennahverkehr zu nutzen. Dabei berücksichtige der Amerikaner die Tatsache, dass die Option, das Auto stehen zu lassen, in normalen Zeiten kaum erwogen werde. Dies ändere sich aber, wenn ein eisiger Wintertag komme. Dann wollen laut Hickel alle Busse und Bahnen verwenden, doch die Kapazitäten reichten nicht. Für die „bessere Annahme des öffentlichen Verkehrsangebots" liefere McFaddan wichtige Hinweise.

In seiner Stellungnahme charakterisiert Hickel die beiden Preisträger als „Forscher in der Stille", deren Untersuchungen der Öffentlichkeit nur schwer zu vermitteln seien. Beide arbeiteten zurückgezogen; selbst an den hiesigen Universitäten würden ihre Erkenntnisse kaum gelehrt. Dennoch gebühre den beiden Amerikanern die Anerkennung, individuelles ökonomisches Verhalten als Basis der gesamtwirtschaftlichen Entwicklung „durchaus praxisrelevant" zu analysieren.

Daniel McFadden

Freude löste die Nachricht aus Stockholm auch in Mannheim beim Zentrum für Europäische Wirtschaftsforschung (ZEW) aus, mit dessen Ressort Arbeitsmärkte, Personalmanagement und Soziale Sicherung Heckmann kooperiert. Er sei für das ZEW bereits in der Vergangenheit als Referent aufgetreten, heißt es in einer Mitteilung des Instituts. Im nächsten Jahr werde er an zwei Konferenzen in Berlin teilnehmen, die er gemeinsam mit dem ZEW organisiere.

Die Begründung der Stockholmer Akademie und Verknüpfungen zu den Webseiten der beiden Preisträger finden sich im weltweiten Datennetz unter der Internetadresse *www.nobel.se*.

Beispiel Mikroökonometrie:

Die Ausgaben privater Haushalte – Ein mikroökonometrisches Modell für die Bundesrepublik Deutschland (Merz 1980)

Mikrotheorie: Theorie des Haushalts, Individuelle Konsumausgaben

Nutzenmaximierung unter Budgetbeschränkung

$$u\left(x_{i1},...,x_{im}\right)=u(x_i)=\max! \qquad s.t.\ \overline{p}_i\cdot\overline{x}_i=y_i$$

Ergebnis: optimale Allokation der Ausgaben auf den Kauf der optimalen Gütermengen

$$v_{ij}=p_{ij}x_{ij}=f\left(\overline{p}_i,y,s\right)$$

v_{ij} = Ausgaben, $\overline{p}_i$ = Preisvektor,

p_{ij} = Preise, y = Einkommen, x_{ij} = Mengen,

s = Sozioökonomische Charakteristika, j = (1,...,m) Gütergruppen,

i = (1,...,n) Individuen (Haushalte)

Komplette Nachfragesysteme: LES, ELES, FELES1, FELES2

LES: Linear Expenditure System (Stone 1954)

$$u(x_{i1},...,x_{im})=\sum_{j=1}^{m}\beta_j\log\left(x_j-\gamma_j\right)=\max! \qquad \text{Stone Geary Nutzenfunktion}$$

s.t. $\overline{p}_i\overline{x}_i=y_i$

Die Nutzenoptimierung unter der Budgetrestriktion führt zu den optimalen Allokationsgleichungen

$$v_{ij}=p_{ij}x_{ij}=\underbrace{p_j\gamma_j}_{\substack{\text{Mindestverbrauch}\\ \text{'basic consumption'}}}+\underbrace{\beta_j}_{\text{marginale Budgetanteile}}\left(y_i-\underbrace{\sum_{k=1}^{m}p_k\gamma_k}_{\substack{\text{supernumerary income}\\ \textit{Mindestausgaben}\\ \text{'}\textit{subsistence}\text{ income'}}}\right)$$

LES-Erweiterungen: Extended Linear Expenditure System (ELES, Lluch 1973) und Functionalized Extended Linear Expenditure System (FELES, Merz 1983)

FELES 1, 2: Functionalized Extended Linear Expenditure System (Merz 1983)

FELES 1: funktionalisierte Minimalausgaben:

$$\gamma_{ij}^{*}=f\left(s,etc.\right)$$

FELES2: funktionalisierte partielle marginale Konsumquoten:

$$\beta_{ij}^{*} = f(s, etc.)$$

Mikrodaten:

EVS '69 mit 47.000 Haushalten (Mikrodatensätzen)

Mikroökonometrie:

Wie Ausgaben modellieren, die

- nicht von jedem Individuum (Haushalt) getätigt werden in einer Periode (Querschnitt) z.B. Kauf dauerhafter Konsumgüter: Auto, Kühlschrank etc.
- bzw. nicht in jedem Jahr eines individuellen Längsschnittes getätigt werden.

Antwort:
Diskrete Entscheidungsmodelle**:** LOGIT, PROBIT, TOBIT, MULTINOMIALES LOGIT etc.

Zweistufiges Ausgabenmodell (Merz 1980)

1. Stufe: Ausgaben dauerhafter Konsumgüter (LOGIT, TOBIT)
2. Stufe: Ausgaben für Verbrauchsgüter (FELES, Systemschätzung)

Neben demographischen und rein ökonomischen Erklärungsfaktoren wurden sozioökonomische Größen zur Erklärung individuellen Verhaltens herangezogen

Insgesamt interessante Ergebnisse: z.T. Sozioökonomische Variablen bedeutender als Preise und Einkommen.

2 Modelle mit diskreten abhängigen Variablen: 'Discrete Choice' und 'Limited Dependent Variables (LDV)'

Diskrete abhängige Variable, qualitativ abhängige Variable

linke Seite einer Gleichung (LHS=left hand side) $y = a + b \cdot x$

- ...ist dichotom: 0,1
- ...ist polytom: 0,1,2,3 ('Zähldaten')

Binäre Entscheidungsmodelle (Kapitel II)	LOGIT, PROBIT etc.
Multiple Entscheidungsmodelle (Kapitel III)	MNL, Nested, 'ordered response', Poisson, simultane Modelle)

Begrenzt abhängige Variable (Limited Dependent Variables, LDV, Kapitel IV)

linke Seite einer Gleichung $y = a + b \cdot x$ ist nur größer (kleiner) als ein Schwellenwert (z.B. $y > 0$). Modelle mit begrenzt abhängigen Variablen (LDV): TOBIT, zweistufiges Heckman-Modell, Modell von Lee ($\Rightarrow$ erst Selektion, dann Ausprägung).

Qualitative Response Models (QRM): Discrete Choice- und LDV-Modelle

Zurückgehend auf Ansätze der Biologie, Medizin gefolgt von der Psychologie und der Soziologie (z.B. Berkson 1944). Anwendungen in der Ökonomie seit den 70er Jahren mit zunehmender Verfügbarkeit von Mikrodaten.

> Zentrale neuere Lehrbuch-Literatur: Greene 2008a,b, Long 1997, Long und Freese 2006, Wooldridge 2006 Chapter 17, Agresti 2007, Winkelmann und Boes 2009, Cameron und Trivedi 2005, Cameron and Trivedi 2010 (with Stata), Gourieroux 2000, Hagenaars 1990

Literatur in deutsch: Tutz 2000, Ronning 1991

Originaltexte: McFadden 1974, 1976, 1985, Mansky und McFadden 1981, Amemiya 1981, Maddala 1983

3 Anwendungsfälle und Beispiele

Charakteristika: Es gibt eine **endliche Anzahl von Alternativen,** die zu erklärenden Entscheidungen sind diskreter Natur

- Kauf oder Nichtkauf: 0,1 (Konsumtheorie)
- Studieren oder nicht
- Arbeitsangebot/Erwerbstätigkeit: Arbeit anbieten oder nicht (0,1)
- Arbeitszeit: 'hours' ≥ 0
- Arbeitslosigkeit: ja/nein, Dauer: positiv stetig (LDV)
- Ausgaben für dauerhafte Konsumgüter: positiv stetig
- Entscheidungen für mehrere Alternativen: polytom (nominal)
- Produktwerbung: polytom, geordnet (ordinal)
- Zeitallokation: Markt- und nichtmarktmäßige Aktivitäten: simultane Entscheidungen und simultane Ausprägung (multiple jobs + multiple hours)
- Wohnungsmarkt, Hausbau (ja/nein)
- Transport: Öffentliche Verkehrsmittel oder nicht
- Renteneintritt: ja/nein
- Gewerkschaftsmitgliedschaft ja/nein
- Aktienkauf/-verkauf
- Portfolioentscheidungen

Dummy dependent variables:

- How do high school students decide whether to go to college?
- What distinguishes PEPSI drinkers from COKE drinkers?

- How can we convince people to commute using public transportation instead of driving?'

Ungeordnete Alternativen: Transportmodi (Auto; y=1; Bus: y=2; Bahn: y=3 …)

Geordnete Alternativen: Einkommensklassen, Marketing Präferenzen

II BINÄRE ENTSCHEIDUNGSMODELLE

1 Struktur binärer Entscheidungsmodelle

y ist eine binäre Variable, die den Wert 1 bei der Wahl einer bestimmten Option oder Eintritt eines Ereignisses oder, wenn nicht, den Wert 0 annimmt. Binäre Entscheidungsmodelle können durch drei Ansätze erklärt werden: als binäres Regresionsmodell, als latentes Regressions-Index Funktionsmodell, oder als stochastisches Nutzenmaximierungsmodell (Random Utility Maximization, RUM). Alle drei Ansätze haben einen gemeinsamen Kern: die Bestimmung der Wahrscheinlichkeit für einen Ereigniseintritt.

1.1 Binärer Regressionsansatz

Analog der Vorgehensweise im klassischen linearen Regressionsmodell wird im binären Regressionsansatz ein ökonometrisches, stochastisches Modell für binäre Entscheidungsmodelle formuliert:

$$y_i = E[y_i] + (y_i - E[y_i])$$
$$y_i = E[y_i] + \varepsilon_i$$

Der Erwartungswert der binären y_i ist

$$\begin{aligned} E[y_i] &= \sum_i y_i \, f(x_i) \\ &= P[y_i = 1] \cdot y_i(=1) + P[y_i = 0] \cdot y_i(=0) \\ &= P[y_i = 1] \cdot 1 \\ &= P(x_i\beta) = F(x_i\beta) \end{aligned}$$

wobei $x_i\beta = \beta_1 x_{i1} + ... + \beta_K x_{iK}$. Damit lautet der allgemeine **binäre Regressionsansatz** für das binäre Entscheidungsmodell

$$E(y_i) = F(x_i\beta) \quad bzw. \quad y_i = F(x_i\beta) + \varepsilon_i$$

wobei $F(x_i\beta)$ eine Wahrscheinlichkeits-Verteilungsfunktion ist (cdf). Je nach konkretem Funktionstyp $F(x_i\beta)$ werden wir alternative binäre Entscheidungsmodelle haben.

1.2 Latentes Regressions-Index Funktionsmodell

In dem Index Funktionsmodell ist der Entscheidungsprozess – das Abwägen der Kosten und Nutzen z.B. ein Auto zu kaufen –, der dann schließlich zu einem beobachtbaren Ereignis führt, nicht beobachtbar, also latent. Wenn ein bestimmter Schwellenwert (‚threshold') – ein Anspruchsniveau, ein Nutzenniveau etc. – dann überschritten ist, ist die Entscheidung gefallen und der Ereigniseintritt beobachtbar.

Latentes Regressions-Index Funktionsmodell

y^* nicht beobachtbar, latent (z.B. Anspruchsniveau für Arbeitsaufnahme)

y beobachtet, eine von zwei Möglichkeiten (gekauft oder nicht, 'work or not')

$x_{(i)}$ Einflussgrößen, Vektor $(x_0,...,x_k)$

β Parametervektor $(\beta_0,...,\beta_k)$ zu schätzen

$\varepsilon_{(i)}$ Störgröße, Residuum

Die latente Variable y^* wird als Regressionsmodell formuliert, überschreitet sie eine Schwelle, können wir das Ereignis $y=1$ beobachten:

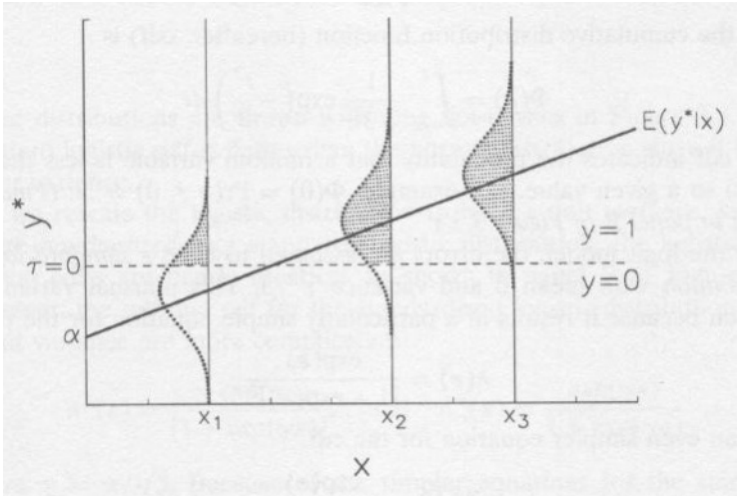
$$y^* = x\beta + \varepsilon$$

$$y = \begin{cases} 1 & \text{falls } y^* > a(=0) \quad \textit{Anspruchs-/Nutzenniveau} \\ 0 & \text{sonst} \end{cases}$$

mit Individualdaten $(i=1,\ldots,n)$:

$$y_i^* = x_i\beta + \varepsilon_i$$

$$y_i = \begin{cases} 1 \text{ falls } & y_i^* > a \\ 0 \text{ sonst} & \end{cases} \quad \text{(a ist Schwellenwert, "threshold")}$$



Latentes Index Funktionsmodell (Long und Freese 1997, S. 41)

Wahrscheinlichkeit für einen Ereigniseintritt

Die Wahrscheinlichkeit für einen Ereigniseintritt aus dem Index Funktionsmodell

$$\begin{aligned} P\left[y_i = 1\right] &= P\left[y_i^* > 0\right] \\ &= P\left[x_i\beta + \varepsilon_i > 0\right] = P\left[\varepsilon_i > -x_i\beta\right] \\ &= 1 - F(-x_i\beta) \\ &= P\left[\varepsilon_i < x_i\beta\right] = F(x_i\beta) \qquad \textit{falls } f(x_i\beta) \textit{ symmetrisch ist.} \end{aligned}$$

$P\left[y_i = 1\right] = P\left[y_i^* > 0\right] = P\left[\varepsilon_i < x_i\beta\right] = F(x_i\beta)$ ist also die Wahrscheinlichkeit dafür, höchstens $x_i\beta$ zu sein, also eine Wahrscheinlichkeits-Verteilungsfunktion $F(x_i\beta)$ (cdf: cumulative distribution function).

Wahrscheinlichkeit dafür, dass das Ereignis *nicht* eintritt:

$$P\left[y_i = 0\right] = 1 - P\left[y_i = 1\right] = 1 - P\left[\varepsilon_i < x_i\beta\right] = 1 - F(x_i\beta) \ .$$

Die 0,1 y-Werte sind Realisationen eines binomialen Prozesses mit der Wahrscheinlichkeit $P\left[y_i = 1\right] = F(x_i\beta)$, die sich von Zug (‚trial') um Zug in Abhängigkeit von x_i ändert.

Beispiel:

y^* = Anspruchsniveau, Nutzenschätzung eines Arbeitsplatzes, propensity to work

Wenn y^* größer als ein Schwellenwert *a* ist (Mindestanforderung an Arbeitsplatz), dann Arbeitsaufnahme ($y = 1$).

1.3 RUM: Random Utility Model

Die nutzentheoretische Fundierung des diskreten Entscheidungsmodells, das RUM-Modell geht auf McFadden 1976, 1985 zurück. Anwendungen im Transportwesen ('urban travel demand': Domencich und McFadden 1975; einen Überblick gibt z.B. Amemiya 1981.

Im stochastischen Nutzenmodell wird jeder diskreten Alternative ein Nutzenindex zugeordnet. Ein rationaler Entscheider wählt nun die Alternative, die ihm den höchsten Nutzen stiftet. Haben wir zwei Alternativen, z.B. ein Haus zu mieten oder es zu kaufen, mit als nicht beobachtbaren Nutzen der Miete und der des Kaufs. Beobachtet wird die Wahl, die den größten Nutzen erbringt, aber nicht die unbeobachtbaren Nutzen. Der entsprechende Indikator (Indexfunktion) ist dann 1 wenn $u_1 > u_2$ und 0, wenn $u_1 \leq u_2$, wenn also die Wahl bspw. auf die Miete fällt.

Stochastische Nutzenspezifikation

$$u_{i1} = x_i\beta_1 + \varepsilon_{i1} \qquad u_{i2} = x_i\beta_2 + \varepsilon_{i2}$$

Wird also die Alternative 1 mit $u_{1i} > u_{2i}$ gewählt, beobachten wir $y_i=1$ und $y_i=0$ sonst und wir erhalten für das binäre Entscheidungsproblem

$$\begin{aligned}
\text{Prob}(y_i = 1 \mid x_i) &= \text{Prob}(u_{1i} > u_{2i}) \\
&= \text{Prob}(x_i\beta_1 + \varepsilon_{i1} - x_i\beta_2 - \varepsilon_{i2} > 0 \mid x_i) \\
&= \text{Prob}(x_i(\beta_1 - \beta_2) + \varepsilon_{i1} - \varepsilon_{i2} > 0 \mid x_i) \\
&= \text{Prob}(x_i\beta + \varepsilon_i > 0 \mid x_i) \\
\\
&= P[\varepsilon_i > -x_i\beta] = 1 - F(-x_i\beta) \\
&= P[\varepsilon_i < x_i\beta] = F(x_i\beta) \qquad \textit{falls } f(x_i\beta) \textit{ symmetrisch ist}
\end{aligned}$$

also den bekannten Ansatz einer binären Regression bzw. des Index Funktionsmodells mit Differenzen in Bezug auf eine Basisalternative (hier Kauf des Hauses).

Wird für die Fehler ε_{i1} und ε_{i2} und damit auch für $\varepsilon_i = \varepsilon_{i1} - \varepsilon_{i2}$ eine Normalverteilung angenommen, dann ergibt sich nach Standardisierung der Varianz auf 1 das PROBIT Modell mit $F(x_i\beta)$ als standardnormale Verteilungsfunktion (cdf).

Wird für ε_i eine Typ I Extremwertverteilung mit $F(x_i\beta) = \exp(-e^{-\varepsilon})$ angenommen, dann ergibt sich das LOGIT Model.

Die Nutzendifferenz ist eine kontinuierliche Zufallsvariable, die auch als latente Größe im Index Funktionsmodell interpretiert werden kann, die, wenn sie über einen Schwellenwert hinausgeht, über die Indexfunktion den Eintritt eines beobachtbaren Ereignisses *j* bestimmt. Die dahinter liegenden Nutzenkomponenten sind nicht direkt beobachtbar.

Damit steht in allen drei Erklärungsansätzen die Wahrscheinlichkeit eines Ereigniseintritts $\text{Prob}(y_i = 1) = F(x_i\beta)$ im Zentrum.

Zusammenfassend kann das binäre Entscheidungsmodell also aus drei Perspektiven entwickelt und erklärt werden: dem

- binären Regressionsmodell,
- Index Funktionsmodell,
- stochastischen Nutzenmodell RUM

mit

Binäres Entscheidungsmodell – Alternative Erklärungsansätze

Binares Regressionsmdell

$$y_i = E[y_i] + \varepsilon_i \qquad y_i = 0,1$$

$$\begin{aligned} E[y_i] &= P[y_i = 1] \cdot y_i(=1) + P[y_i = 0] \cdot y_i(=0) \\ &= P[y_i = 1] \cdot 1 \\ &= P(x_i\beta) = F(x_i\beta) \end{aligned}$$

$$y_i = F(x_i\beta) + \varepsilon_i \qquad y_i = 0,1$$

Index Funktionsmodell

$y_i^* = x_i\beta + \varepsilon_i$ $\qquad y_i^* = latent$ $\qquad \varepsilon_i$ ist Zufallsterm $(E[\varepsilon_i] = 0)$

$$y_i = \begin{cases} 1 \text{ falls } & y_i^* > a \\ 0 \text{ sonst} & \end{cases}$$ (a als Anspruchsniveau, "threshold")

Wahrscheinlichkeit für den Ereigniseintritt

$$\begin{aligned} P[y_i = 1] &= P[y_i^* > 0] \\ &= P[x_i\beta + \varepsilon_i > 0] = P[\varepsilon_i > -x_i\beta] = 1 - F(-x_i\beta) \\ &= P[\varepsilon_i < x_i\beta] = F(x_i\beta) \qquad \text{falls } f(x_i\beta) \text{ symmetrisch ist.} \end{aligned}$$

RUM: Random Utility Model

$$u_{i1} = x_i\beta_1 + \varepsilon_{i1} \qquad u_{i2} = x_i\beta_2 + \varepsilon_{i2}$$

$$\begin{aligned} \text{Prob}(y_i = 1 \mid x_i) &= \text{Prob}(u_{1i} > u_{2i}) \\ &= \text{Prob}(x_i\beta_1 + \varepsilon_{i1} - x_i\beta_2 - \varepsilon_{i2} > 0 \mid x_i) \\ &= \text{Prob}(x_i(\beta_1 - \beta_2) + \varepsilon_{i1} - \varepsilon_{i2} > 0 \mid x_i) \\ &= \text{Prob}(x_i\beta + \varepsilon_i > 0 \mid x_i) \end{aligned}$$

$$\begin{aligned} &= P[\varepsilon_i > -x_i\beta] = 1 - F(-x_i\beta) \\ &= P[\varepsilon_i < x_i\beta] = F(x_i\beta) \qquad \text{falls } f(x_i\beta) \text{ symmetrisch ist} \end{aligned}$$

2 Alternative Modelle

2.1 Das lineare Wahrscheinlichkeitsmodell: LP

Die wohl einfachste Annahme ist, für F die Identitätsfunktion I anzunehmen: $F(x_i\beta) = I(x_i\beta) = x_i\beta$. Damit ergibt sich das Lineare Wahrscheinlichkeitsmodell LP mit

$$y_i = F(x_i\beta) + \varepsilon_i = I(x_i\beta) + \varepsilon_i = x_i\beta + \varepsilon_i \qquad \text{mit } \varepsilon_i \text{ als Zufallsvariable und } E[\varepsilon_i] = 0$$

Das lineare Wahrscheinlichkeitsmodell entspricht damit dem bekannten klassischen linearen Regressionsansatz mit $\varepsilon_i \sim N(0,\sigma^2)$; es wäre schön, alle Resultate davon übertragen zu können.

Aber:

1. Verteilung von $(\varepsilon_i \mid x_i)$ ist nicht 'normal'

Bei geg. x_i hat ε_i keine stetige und normale, sondern eine diskrete Wahrscheinlichkeitsfunktion mit nur zwei Ausprägungen; da $y_i = x_i\beta + \varepsilon_i$ nur zwei Werte (0,1) annehmen kann, kann ε_i auch nur zwei Werte annehmen:

$$y_i = x_i\beta + \varepsilon_i \quad mit$$
$$1_i = x_i\beta + (1 - x_i\beta) \quad oder$$
$$0_i = x_i\beta + (-x_i\beta)$$

y_i	ε_i	$P(\varepsilon_i)$
1	$1 - x_i\beta$	$P_i = x_i\beta$
0	$-x_i\beta$	$1 - P_i = 1 - x_i\beta$

z.B.: ist für $y_i(0,1) = x_i\beta + \varepsilon_i$ der deterministische Teil $x_i\beta = 13$, dann ist für
$y(=1) = 13 + (1-13)$ der Fehlerterm $\varepsilon_i = 1 - 13$ und für
$y(=0) = 13 + (-13)$ der Fehlerterm $\varepsilon_i = -13$, also nur zweiwertig möglich.

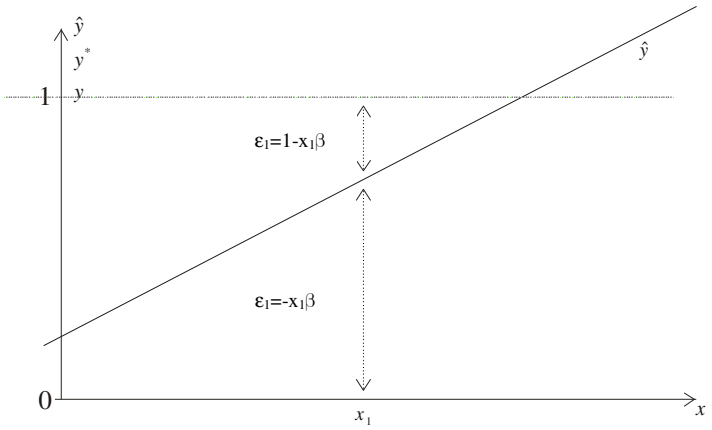
Daraus folgt: Konventionelle Standardfehler, t-Werte, R^2 etc. sind verzerrt.

2. ‚Wahrscheinlichkeit' kann außerhalb des Intervalls $(0 \le P \le 1)$ liegen

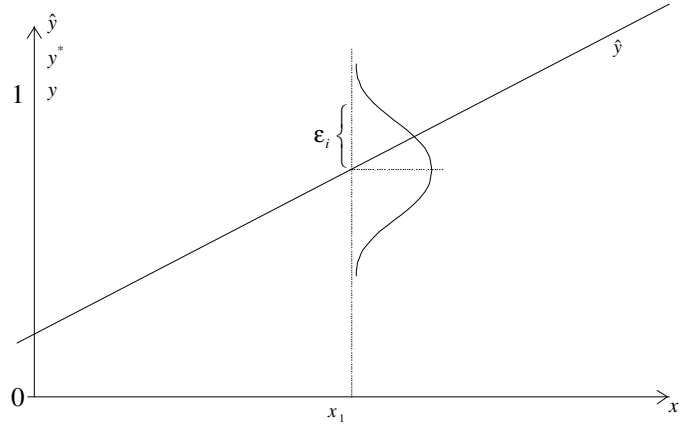
$$\begin{aligned} E[\varepsilon_i] &= \sum_{\varepsilon_i} \varepsilon_i f(\varepsilon_i) \\ &= (1 - x_i\beta) P_i + (-x_i\beta)(1 - P_i) \\ &= P_i - x_i\beta P_i - x_i\beta + x_i\beta P_i \\ &= P_i - x_i\beta = 0 \end{aligned}$$

$E[\varepsilon_i] = 0$ also nur dann, wenn $P_i = x_i\beta$. Wenn also nicht $0 \le x_i\beta \le 1$, dann wäre eine Wahrscheinlichkeit kleiner Null oder größer 1 gegeben, was aber offensichtlich der Intention $0 \le P_i \le 1$ widerspricht.

ε_i ist dichotom:

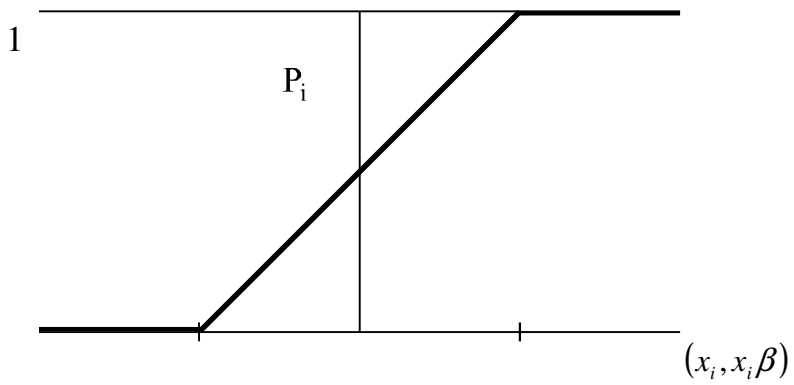


und nicht stetig wie im klassischen linearen Regressionsmodell (CLR):



3. Vorhersagen $\hat{P}_i = x_i\hat{\beta}$ können ebenso außerhalb des Intervalls [0,1] liegen.

Vorhersagen $\hat{P}_i = x_i\hat{\beta}$ können ebenso außerhalb des Intervalls [0,1] liegen wegen Beschränkung auf [0,1] mit $\hat{p}_i = \begin{cases} 1 & \text{wenn } x_i\hat{\beta} > 1 \\ 0 & \text{wenn } x_i\hat{\beta} < 0 \end{cases}$; führt zu 'Knicken' und Verzerrungen.



4. Störterme sind heteroskedastisch.

$$\varepsilon_i = \begin{cases} 1 - x_i\beta & \text{x}_\text{i} \text{ veränderlich} \\ -x_i\beta & \end{cases}$$

Allgemein ist die Varianz von ε_i

$$\begin{aligned} Var[\varepsilon_i] &= E\left[\left(\varepsilon_i - E[\varepsilon_i]\right)^2\right] \\ &= \sum_{\varepsilon_i}\left(\varepsilon_i - E[\varepsilon_i]\right)^2 f(\varepsilon_i) \qquad \varepsilon_i \text{ diskret} \\ &= \sum_{\varepsilon_i}(\varepsilon_i)^2 f(\varepsilon_i) \qquad E[\varepsilon_i] = 0 \end{aligned}$$

Im linearen Wahrscheinlichkeitsmodell LP ist die Varianz von ε_i mit den zwei Ausprägungen von ε_i

$$\begin{aligned} spez.\,Var[\varepsilon_i] &= \sum_{j=1}^{2}(\varepsilon_{ji} - 0)^2 f(\varepsilon_{ji}) = \varepsilon_{1i}^2 P_i + \varepsilon_{2i}^2(1 - P_i) \\ &= (1 - x_i\beta)^2 \cdot P_i + (-x_i\beta)^2 \cdot (1 - P_i) \\ &= (1 - x_i\beta)^2 \cdot x_i\beta + (x_i\beta)^2 \cdot (1 - x_i\beta) \\ &= \left[(1 - x_i\beta)^2 + (x_i\beta)\cdot(1 - x_i\beta)\right] x_i\beta \\ &= \left[1 - 2x_i\beta + (x_i\beta)^2 + x_i\beta - (x_i\beta)^2\right] x_i\beta \\ &= (1 - x_i\beta)\, x_i\beta \\ &= (1 - P_i)\cdot P_i \end{aligned}$$

damit ist wegen $\varepsilon_i = f(x_i)$ mit seiner Abhängigkeit von jeder Beobachtung i die $\text{var}(\varepsilon_i)$ nicht gleich bleibend für alle i, homoskedastisch, sondern veränderlich mit x_i

$\Rightarrow$ LP ist heteroskedastisch; damit sind die OLS (MKQ)-Resultate nicht effizient.

Insgesamt: das lineare Wahrscheinlichkeitsmodell LP ist also nicht adäquat.

Lineares Wahrscheinlichkeitsmodell LP

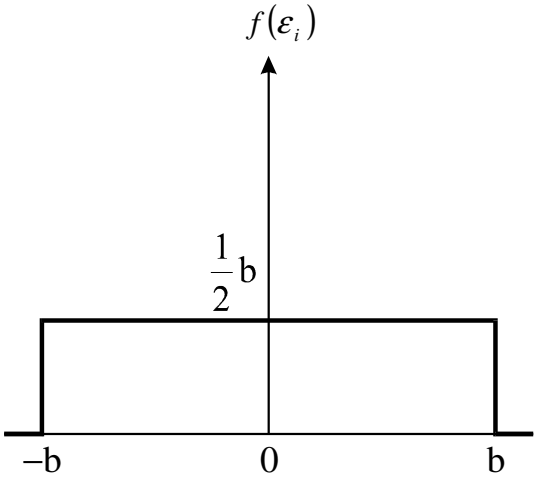
$$y_i = P[y_i = 1] + \varepsilon_i = F(x_i\beta) + \varepsilon_i$$

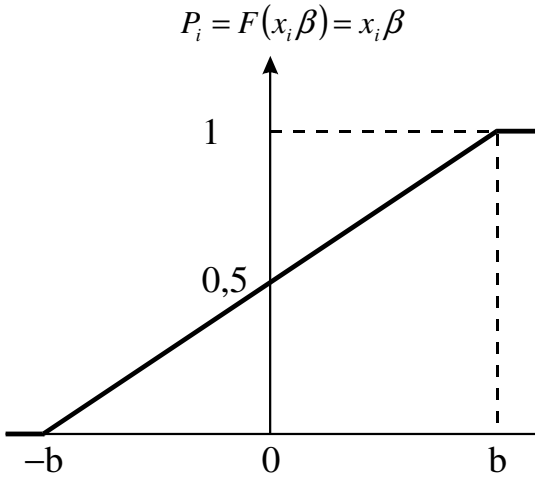
$$P[y_i = 1] = \begin{cases} 0 & wenn\ x_i\beta \le 0 \\ I(x_i\beta) = x_i\beta & wenn\ 0 < x_i\beta < 1 \\ 1 & wenn\ x_i\beta \ge 1 \end{cases}$$

$$y_i = x_i\beta + \varepsilon_i$$

Dichte- und Verteilungsfunktion der Gleichverteilung

mit gleichverteilter Dichtefunktion und linearer Verteilungsfunktion:





Schätzung: zweistufiger 'weighted least squares' (WLS) Ansatz (effizienter als OLS)

1. Stufe $\quad OLS \rightarrow P_i \rightarrow Schätzung\ \hat{\sigma}_i = \sqrt{\hat{P}_i\left(1-\hat{P}_i\right)}$

2. Stufe $\quad OLS\ von\ \left(y_i / \hat{\sigma}_i\right) = \left(x_i / \hat{\sigma}_i\right)\beta + \varepsilon_i^*$

Zwar asymptotisch effizient aber:

- was ist zu tun, wenn $\hat{P}_i$ außerhalb [0,1]? $(\hat{\sigma}_i = 0)$
- sensitiver in der Spezifikation als OLS

$\Rightarrow$ auch durch modifizierte Schätzung: die Schwächen in der Spezifikation sind nicht behebbar.

Für eine ausführliche Diskussion und Reinterpretation des linearen Wahrscheinlichkeitsmodells, vgl. Aldrich, J. und F. Nelson 1984, Linear Probability, Logit and Probit Models, Beverly Hills, Calif.: Sage

2.2 Nichtlineare Transformation I: LOGIT

Allgemeiner Ansatz:

$$y_i = F(x_i\beta) + \varepsilon_i \qquad y_i = 0,1$$

Im linearen Wahrscheinlichkeitsmodell war $F(x_i\beta) = I(x_i\beta) = x_i\beta \Leftrightarrow y_i = x_i\beta + \varepsilon_i$.

Frage: wie sieht das Modell bei anderen Verteilungen $F(x_i\beta) = P[y_i = 1]$ aus?

Da die Verteilung von ε_i im linearen Wahrscheinlichkeitsmodell zu nicht wünschenswerten Ergebnissen führt, werden für ε_i und damit auch für $F(x_i\beta)$ andere Verteilungen angenommen; besonders bedeutend sind die logistische Verteilung und die Normalverteilung:

- logistisch verteilt $\Rightarrow$ LOGIT
- normalverteilt $\Rightarrow$ PROBIT.

LOGIT-Modell: Logistische Verteilung

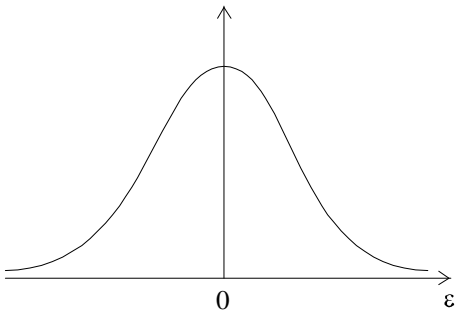
Dichtefunktion (pdf) $$f(\varepsilon_i) = \frac{\mu \cdot e^{-\mu\cdot\varepsilon_i}}{\left(1+e^{-\mu\cdot\varepsilon_i}\right)^2}$$

Verteilungsfunktion (cdf) $$F(\varepsilon_i) = \Lambda(\varepsilon_i) = \frac{1}{1+e^{-\mu\cdot\varepsilon_i}} = \frac{e^{\mu\cdot\varepsilon_i}}{1+e^{\mu\cdot\varepsilon_i}} \qquad \mu>0, -\infty<\varepsilon_i<\infty$$

Logistische Dichte- und Verteilungsfunktion

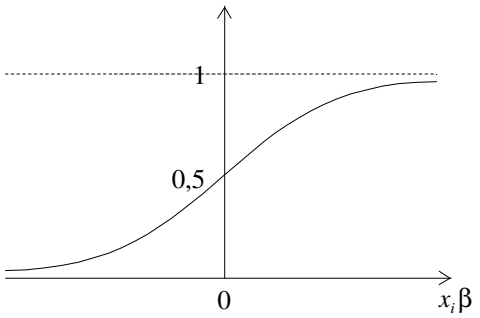
Logistische Dichtefunktion (pdf)

$$f(\varepsilon) = \frac{e^{-\varepsilon_i}}{\left(1+e^{-\varepsilon_i}\right)^2}$$



Logistische Verteilungsfunktion (cdf)

$$F(x_i\beta) = 1/\left(1+e^{-x_i\beta}\right)$$



Ist nun im allgemeinen Ansatz des 'latenten Modells'

$$y_i = \begin{cases} 1 & \text{falls } y_i^* = x_i\beta + \varepsilon > a \\ 0 & \text{sonst} \end{cases}$$

ε logistisch verteilt mit $E(\varepsilon) = 0$ und $Var(\varepsilon) = \pi^2/3$, dann ist auch die Zufallsvariable $y_i^* = x_i\beta + \varepsilon_i$ logistisch verteilt, und mit

$$E[y_i] = P[y_i = 1]\, y_i(=1) + P[y_i = 0]\, y_i(=0) = P[y_i = 1]1 = F(x_i\beta) \text{ ist}$$

$$\begin{aligned} P[y_i = 1] &= P[\varepsilon < x_i\beta] \\ &= \frac{e^{x_i\beta}}{1+e^{x_i\beta}} = \frac{\exp(x_i\beta)}{1+\exp(x_i\beta)} \\ &= \frac{1}{1+e^{-x_i\cdot\beta}} = \Lambda(x_i\beta) \qquad \text{Normierung allg. mit } \mu=1 \end{aligned}$$

Simple Umformungen ergibt die Beziehung zwischen Argument $(x_i\beta)$ und Funktion $\Lambda(x_i\beta)$

$$\begin{aligned} & P[y_i = 1] = \frac{e^{x_i\beta}}{1+e^{x_i\beta}} \\ & P_i\left(1+e^{x_i\beta}\right) = e^{x_i\beta} \\ & P_i = e^{x_i\beta} - P_i e^{x_i\beta} = e^{x_i\beta}(1-P_i) \\ & \frac{P_i}{1-P_i} = e^{x_i\beta} \\ \rightarrow \quad & \log\frac{P_i}{1-P_i} = x_i\beta \qquad \text{log=logarithmus naturalis (ln)} \end{aligned}$$

LOGIT Transformation

$$P[y_i = 1] = \Lambda(x_i\beta) = \frac{e^{x_i\beta}}{1+e^{x_i\beta}} = \frac{\exp(x_i\beta)}{1+\exp(x_i\beta)} = \frac{1}{1+e^{-x_i\cdot\beta}}$$

$$\text{logit}(P_i) = \Lambda^{-1}(P_i) = \log\frac{P_i}{1-P_i} = x_i\beta$$

Damit ergibt sich ein Schätzansatz

$$\log\frac{P_i}{1-P_i} = x_i\beta + \varepsilon_i \qquad \left[\varepsilon \quad \text{z.B. } N\left(0,\sigma_\varepsilon^2\right)\right],$$

der sich im Gegensatz zum linearen Wahrscheinlichkeitsmodell dort mit $y_i = (1 \; oder \; 0) = x_i\beta + \varepsilon$ (dort ε eben nicht $N\left(0,\sigma_\varepsilon^2\right)$) jetzt adäquater mit OLS schätzen lässt.

„LOGIT and PROBIT cannot be estimated using the usual least squares and weighted least squares techniques… Without grouping the data, we can not possibly form the empirical probabilities as the number of successes divided by the number of trials. Each observation represents 0

and 1 outcome of a single (Bernoulli) trial. Nor can we construct the empirical logits and probits from the sample data“ (Powers und Xie (2000, S. 61).

Odds: Das Chancenverhältnis O= P/(1-P) wird auch 'odds' genannt. Odds (O) sind **relative Risiken** und werden zur Darstellung von Häufigkeiten (frequencies) resp. Wahrscheinlichkeiten in log-linearen Modellen oft verwendet und sind allgemein das Verhältnis der Häufigkeit einer Variablenausprägung zu einer anderen.

Interpretation: Bei P=0,80% korrespondiert eine Chance von 4 (0,8/0,2), dass das Ereignis y=1 eintritt; P für y=1 ist viermal höher als für den Nichteintritt. Odds entsprechen Wettquoten.

Wegen $P(y=1)=e^{x_i\beta}/(1+e^{x_i\beta})$; $P(1+e^{x_i\beta})=e^{x_i\beta}$; $P=e^{x_i\beta}-Pe^{x_i\beta}$; $P=e^{x_i\beta}(1-P)$ gilt

$$\frac{\mathrm{P(y=1)}}{\mathrm{1\text{-}P(y=1)}}=\mathrm{e}^{\mathrm{x_i}\beta}.$$

Der Odds-Faktor einer erklärenden Variablen ist der antilog zu β_k, also e^{β_k}; d.h. bei einer Änderung von x_k um δ Einheiten verändern sich die Odds (der linken Seite) um den Faktor $e^{\beta_k\delta}$. Wäre $\beta_k=0{,}1$, dann wird mit einer Erhöhung von x_k um eine Einheit das ursptüngliche Odds-Verhältnis ehöht durch die Multiplikation mit $e^{0{,}1}\simeq 1{,}105$. Das entspricht einer proportionalen Erhöhung von 0,105 der ursprünglichen Odds; die relative Wahrscheinlichkeit des $y=1$ Ereignisses hat sich um 10,5 % erhöht.

β_k ist zugleich eine Semi-Elastizität: für sehr kleine β_k ist $e^{\beta_k}-1\simeq\beta_k$.

Odds ratio: Die odds ratio misst den Einfluss auf die odds einer Steigerung um eine Einheit in nur einer unabhängigen Variablen

$$\text{Odds ratio}=\text{odds}_1/\text{odds}_0$$

Mit odds_1 bei gegebener Steigerung, odds_0 vor der Steigerung. Die odds ratio gibt also eine prozentuale Änderung der odds bezogen auf die Ausgangssituation an.

Sei x_j eine dummy Variable, dann gilt

$$odds\ ratio=e^{\beta_j x_j}$$
$$>1=\ positiver\ Effekt$$
$$<1=negativer\ Effekt$$

Mehr zur Interpretation der Parameter log-linearer Modelle bspw. in Alba 1988, Long 1997, 79ff, Anderson et al. 2010 oder Cameron und Trivedi 2005, 470.

Beispiel Odds

Wenn 100 Personen einen bestimmten Autotyp kaufen und 25 ihn nicht kaufen, dann ist mit O=100/25=4 die Anzahl der Käufer viermal so groß wie die der Nichtkäufer.

Mit P=100/125=0,80 sind von allen 125 hergestellten Autos 80% verkauft.

Würde eine Schätzgleichung anzeigen, dass die Wahrscheinlichkeit eines Autokaufs aus

$$\log\left(\frac{P_i(Kauf)}{1-P_i(Kauf)}\right)=\ldots+0,693\,SEX\ldots$$

vom Geschlecht des Käufers (SEX=1 für Mann, =0 für Frau) abhängt, dann ist die Kaufwahrscheinlichkeit von Männern auf Grund des positiven Vorzeichens des Koeffizienten höher als die von Frauen, bzw. mit $e^{0,693}=2$ sind die Kaufodds eines Mannes ceteris paribus doppelt so hoch wie die einer Frau.

Ein Schätzansatz über die ‚odds' mit $P_i/(1-P_i)=x_i\beta+\varepsilon_i$ könnte gewählt werden ist aber noch restriktiv, da über die odds die Wahrscheinlichkeit P_i nur auf das Intervall $[0,\infty)$ transformiert wird.

$\log\left(P/(1-P)\right)$ allerdings transformiert auf $(-\infty,+\infty)$; damit ist das klassische linear Regressionsmodell mit OLS auf das log-lineare Modell anwendbar.

Logistisches Modell:

$$P\left(x_i\beta\right)=F\left(x_i\beta\right)=\Lambda\left(x_i\beta\right)$$

$$F^{-1}\left(P_i\right)=x_i\beta \;\; mit \;\; F^{-1}\left(P_i\right)=\Lambda^{-1}\left(P_i\right)=\log\frac{P_i}{1-P_i}$$

Das LOGIT-Modell ist deshalb so beliebt, da es für $F^{-1}\left(P_i\right)=\Lambda^{-1}(P_i)=\mathrm{x}_i\beta$ eine geschlossene Lösung und zudem in sehr einfacher Form, nämlich $x_i\beta=\log\frac{P_i}{1-P_i}$, gibt. Für das PROBIT-Modell wird dafür ein Integral $F^{-1}(P_i)$ zu lösen sein.

LOGIT Modell

$$y_i = P[y_i = 1] + \varepsilon_i = F(x_i\beta) + \varepsilon_i$$

$$P[y_i = 1] = \Lambda(x_i\beta) = \frac{1}{1+e^{-x_i \cdot \beta}} \qquad F^{-1}(P_i) = \Lambda^{-1}(P_i) = \log\frac{P_i}{1-P_i} = x_i\beta$$

$$logit(P_i) = \log\frac{P_i}{1-P_i} = x_i\beta$$

$$y_i = \Lambda(x_i\beta) + \varepsilon_i$$

2.3 Nichtlineare Transformation II: PROBIT

Wird statt der logistischen Verteilung die Normalverteilung für ε_i angenommen, mit $E[\varepsilon] = \mu$ und $\mathrm{Var}[\varepsilon] = \sigma^2$, also $\varepsilon_i \sim N(\mu, \sigma^2)$, dann ist die Zufallsvariable $y_i^* = x_i\beta + \varepsilon_i$ normalverteilt, und mit

$$E[y_i] = P[y_i = 1]\, y_i(=1) + P[y_i = 0]\, y_i(=0) = F(x_i\beta)$$

ist $P_i = F(x_i\beta)$ normalverteilt mit

$$P[y_i = 1] = \int_{-\infty}^{x_i\beta} \frac{e^{-t^2/2}}{\sqrt{2\pi}}\, dt = \Phi(x_i\beta)$$

Standardnormalverteilung mit

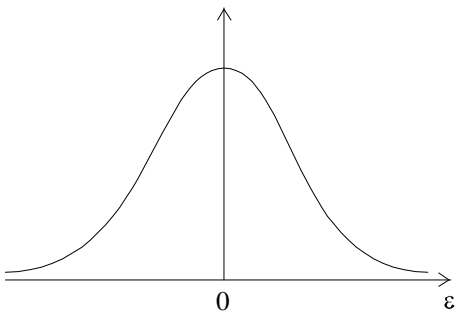
$$z = \frac{x_i\beta - \mu}{\sigma}, \quad \mathrm{N}(0,1)$$

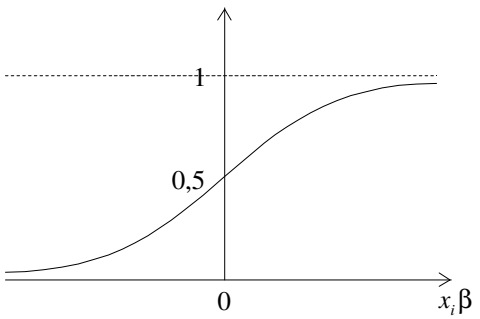
$$E(z) = 0, \quad Var(z) = 1$$

Dichte- und Verteilungsfunktion der Standard-Normalverteilung

(standard-) normalverteilte Dichtefunktion (standard-) normalverteilte Verteilungsfunktion

$$\phi(\varepsilon) = (2\pi)^{-1/2} \exp(\varepsilon^2/2) \qquad \Phi(x_i\beta/\sigma) = \int_{-\infty}^{x_i\beta/\sigma} \phi(t)dt$$





Aus der Standardisierung folgt, dass nicht mehr β später geschätzt werden kann, sondern nur noch β/σ.

PROBIT Transformation

$$P\left[y_i = 1\right] = \Phi(\mathrm{x}_i\beta) = \int_{-\infty}^{x_i\beta} \frac{e^{-t^2/2}}{\sqrt{2\pi}}\, dt$$

$$probit(P) = \Phi^{-1}(P_i) = x_i\beta$$

Wegen der zugrundeliegenden Normalverteilung wird das PROBIT Modell gelegentlich auch als NORMIT Modell bezeichnet.

Durch das Integral werden die notwendigen Berechnungen aufwendiger:

Nun keine simple Umformung für $F^{-1}(P_i)$, so dass

$$F^{-1}(P_i) = \Phi^{-1}\left(P_i\right) = x_i\beta$$

einfach zu berechnen und zu schätzen wäre.

Notwendig hierzu ist es, für jedes i mit $x_i\beta$ das Integral nach $x_i\beta$ aufzulösen (für das LOGIT Modell kann man ja $x_i\beta$ direkt ausrechnen mit $F^{-1}\left(P_i\right) = \log\left(P_i/\left(1-P_i\right)\right) = x_i\beta$

PROBIT Modell

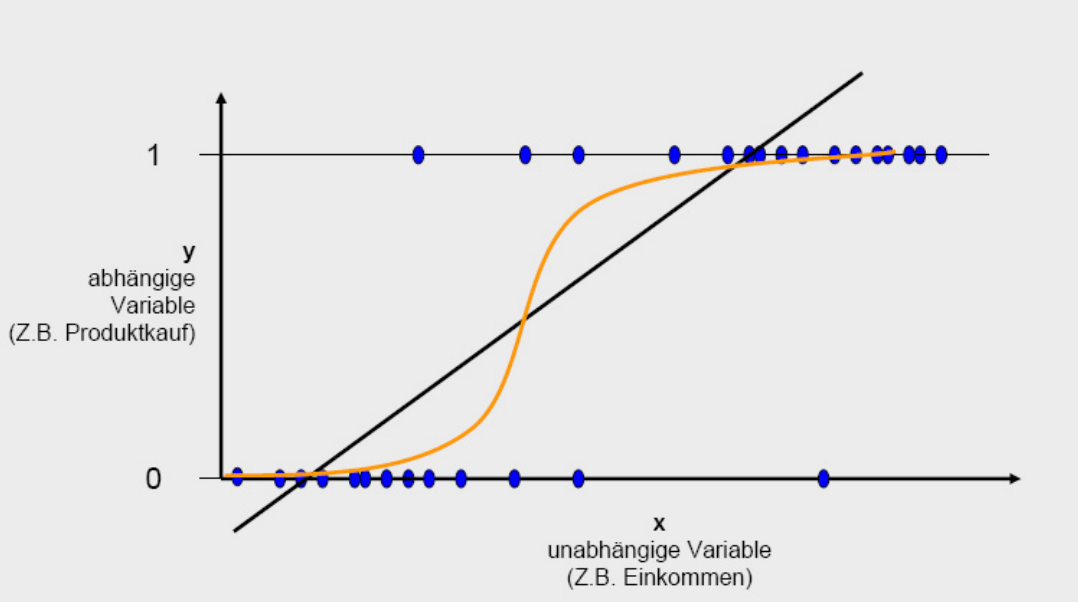
$$y_i = P[y_i = 1] + \varepsilon_i = F(x_i\beta) + \varepsilon_i$$

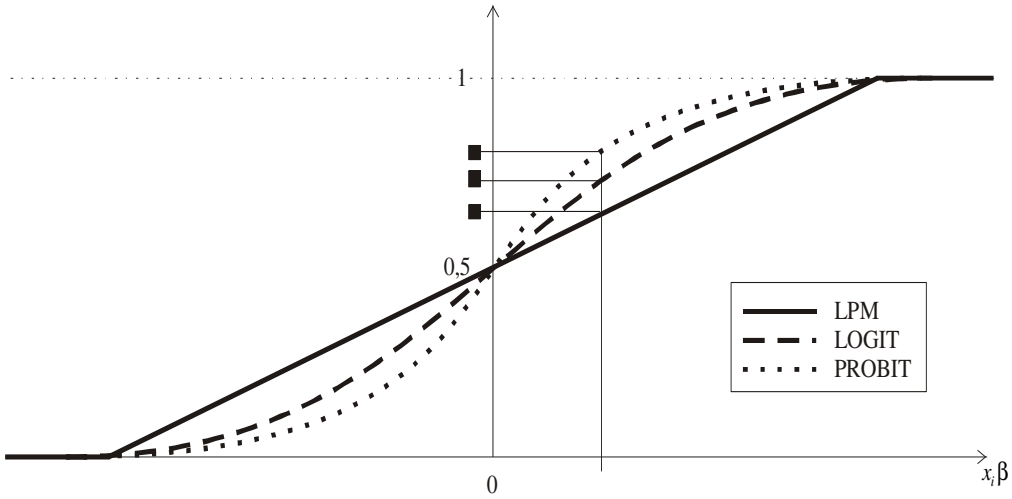
$$P[y_i = 1] = \Phi\left(x_i\beta\right) = \int_{-\infty}^{x_i\beta} \frac{e^{-t^2/2}}{\sqrt{2\pi}}\, dt \qquad F^{-1}\left(P_i\right) = \Phi^{-1}(P_i) = x_i\beta$$

$$probit(P_i) = \Phi^{-1}(P_i) = x_i\beta$$

$$y_i = \Phi(x_i\beta) + \varepsilon_i$$

Beispiel Alternative Anpassungen an binäre Entscheidungen:





Um $P = 0{,}5$ sind alle drei Ansätze – LPM, LOGIT und PROBIT – ähnlich. Für kleine P (nahe Null) und große Wahrscheinlichkeiten (P nahe Eins) werden die Unterschiede zwischen den Modellen allerdings größer.

2.4 Weitere binäre Entscheidungsmodelle: Arctan- gestutzt exponentielle Modelle und weitere Verteilungen

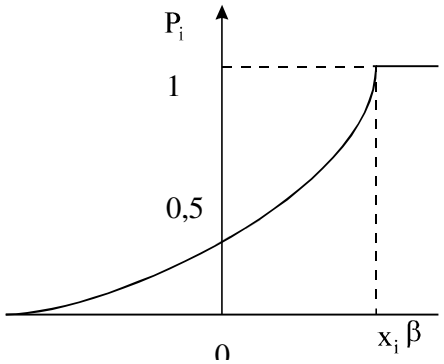
Wenig gebräuchlich, aber nennenswert (mit wünschenswerten Eigenschaften: Ben Akiva und Lerman 1985, S.73)

1. Arctan-Wahrscheinlichkeitsmodell

$$P_i = \frac{1}{2} + \frac{1}{\pi} \cdot \tan^{-1}(x_i \beta)$$

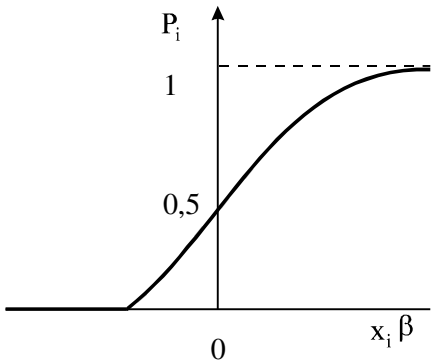
2. Rechts-gestutzt exponentiell (right truncated exponential)

$$P_i = \begin{cases} e^{-(a - x_i\beta)} & x_i\beta < a \\ 1 & \text{sonst} \end{cases}$$



3. Links-gestutzt exponentiell

$$P_i = \begin{cases} e^{-(a - x_i\beta)} & x_i\beta > a \\ 1 & \text{sonst} \end{cases}$$



oder bspw. das GOMPIT Modell (Zellner und Lee (1965) mit

4. GOMPIT Transformation

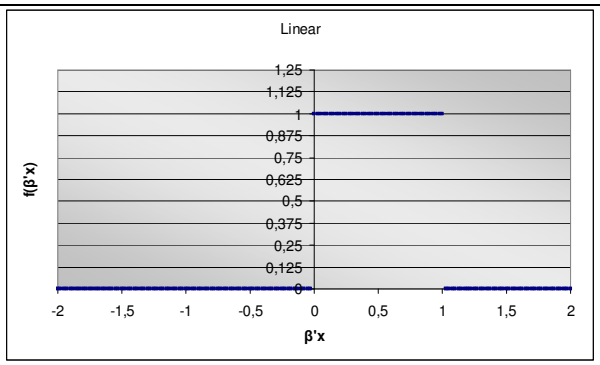
$$P[y_i = 1] = F(x_i\beta) = e^{-e^{-x_i\beta}}$$

$$gompit(P_i) = \log\log(1/P_i) = x_i\beta$$

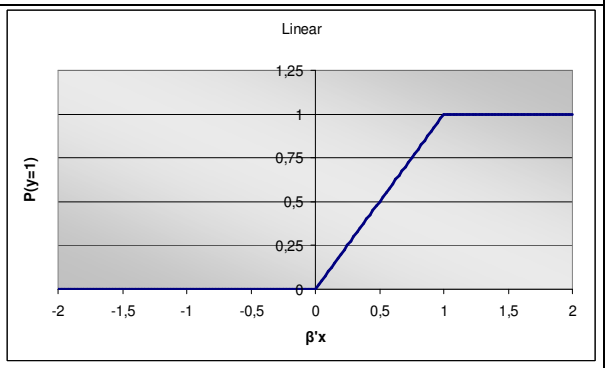
Prominent sind aber LOGIT und PROBIT.

Verteilungs- und Dichtefunktionen für Binary Choice-Modelle

Lineare Dichte- und Verteilungsfunktion

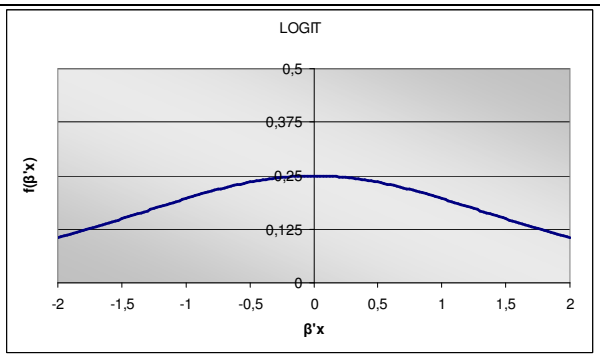


$$f(x_i\beta)=\begin{cases}1 & \text{für } 0<x_i\beta\le 1\\ 0 & \text{sonst}\end{cases}$$

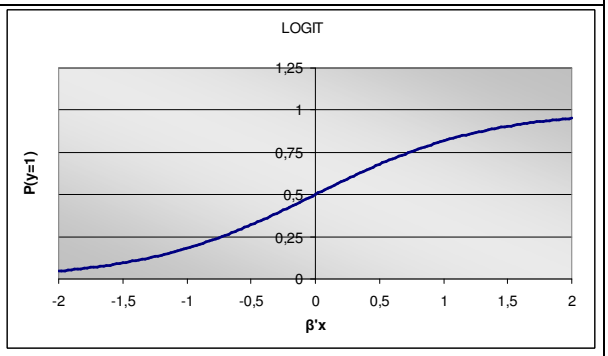


$$P(y_i=1)=\begin{cases}0 & \text{für } x_i\beta\le 0\\ x_i\beta & \text{für } 0<x_i\beta\le 1\\ 1 & \text{sonst}\end{cases}$$

Logistische Dichte- und Verteilungsfunktion

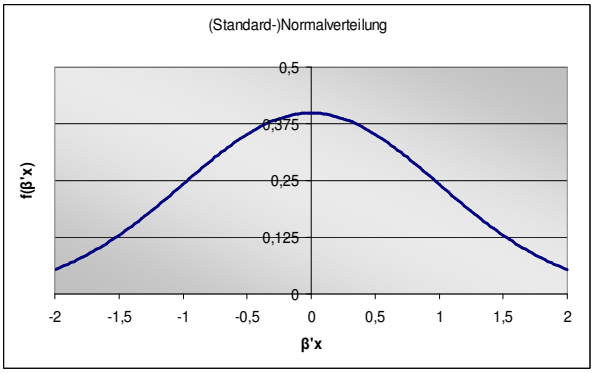


$$f(x_i\beta)=\Lambda(x_i\beta)\left(1-\Lambda(x_i\beta)\right)$$

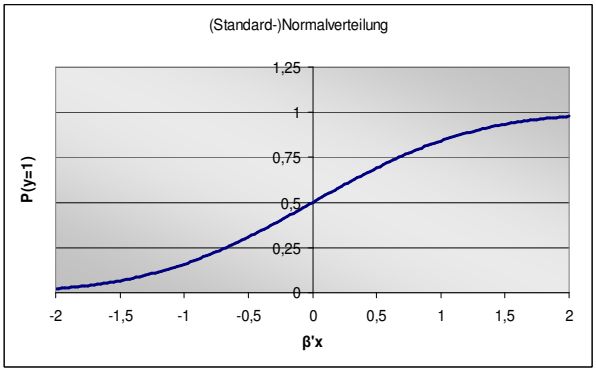


$$P(y_i=1)=\frac{e^{\beta' x_i}}{1+e^{\beta' x_i}}=\Lambda(x_i\beta)$$

Standardnormalverteilte Dichte- und Verteilungsfunktion

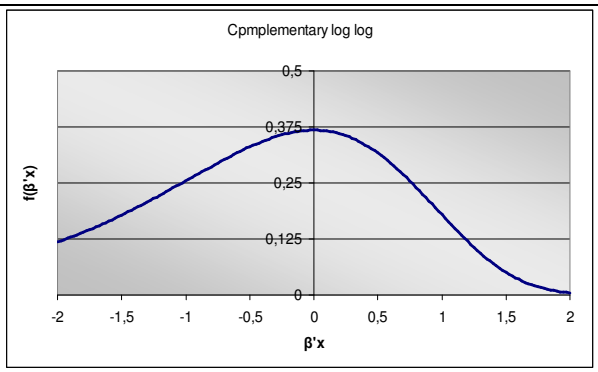


$$f(x_i\beta)=\phi(x_i\beta)$$

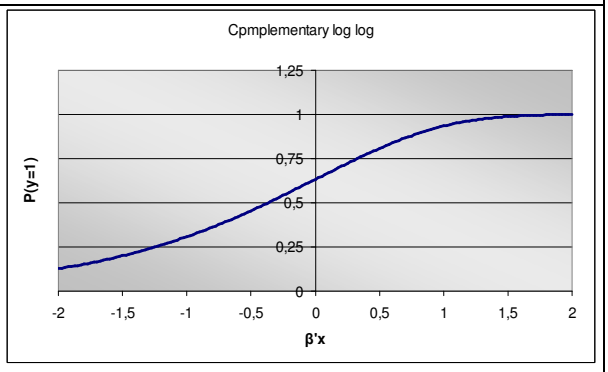


$$P(y_i=1)=\int_{-\infty}^{x_i\beta}\frac{e^{-t^2/2}}{\sqrt{2\pi}}dt=\Phi(x_i\beta)$$

Complemetary log log (Extreme Value) Dichte- und Verteilungsfunktion

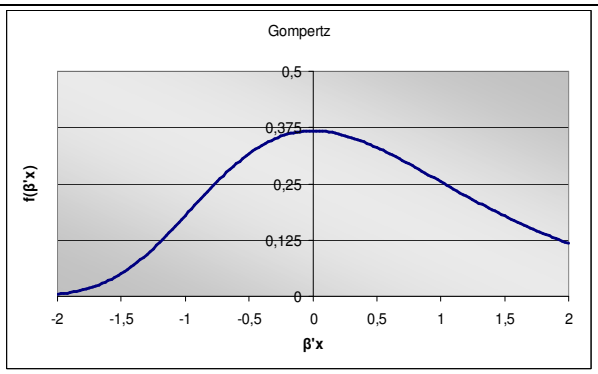


$$f(\mathrm{x_i}\beta) = e^{\beta'\mathrm{x_i}} \cdot \left(1 - C(\mathrm{x_i}\beta)\right)$$

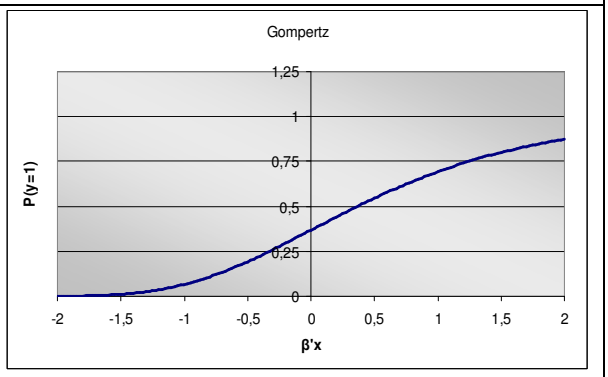


$$P(y_i = 1) = 1 - e^{-e^{\mathrm{x_i}\beta}} = C(\mathrm{x_i}\beta)$$

Gompertz (Gompit) Dichte- und Verteilungsfunktion

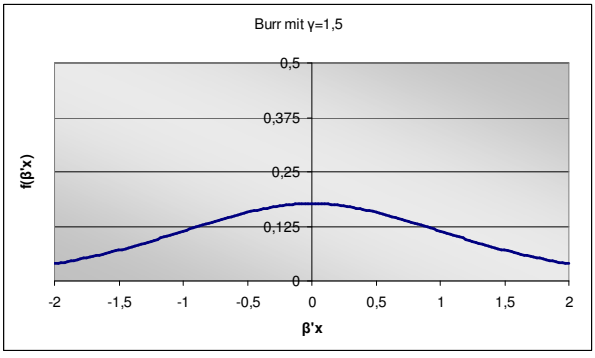


$$f(\mathrm{x_i}\beta) = e^{-\mathrm{x_i}\beta} G(\mathrm{x_i}\beta)$$

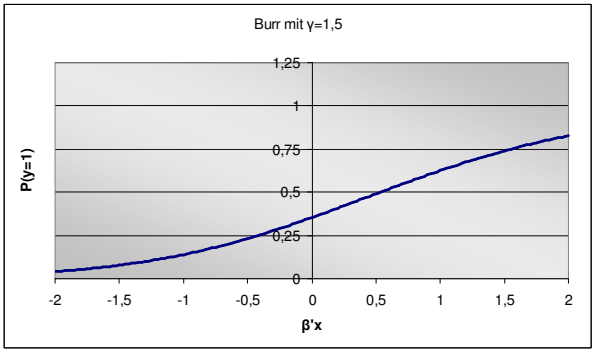


$$P(y_i = 1) = e^{-e^{-\mathrm{x_i}\beta}} = G(\mathrm{x_i}\beta)$$

Burr (Scobit) Dichte- und Verteilungsfunktion

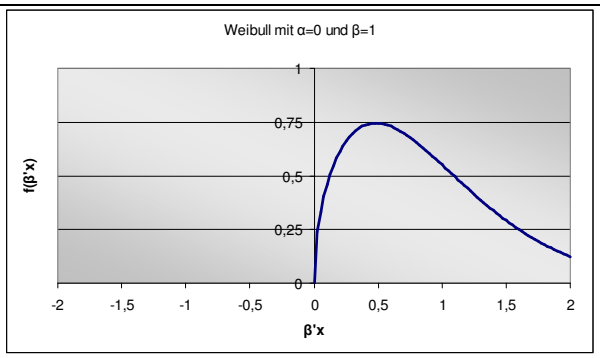


$$f(\mathrm{x_i}\beta) = \gamma\left(\Lambda(\mathrm{x_i}\beta)\right)^{\gamma}\left(1 - \Lambda(\mathrm{x_i}\beta)\right)$$

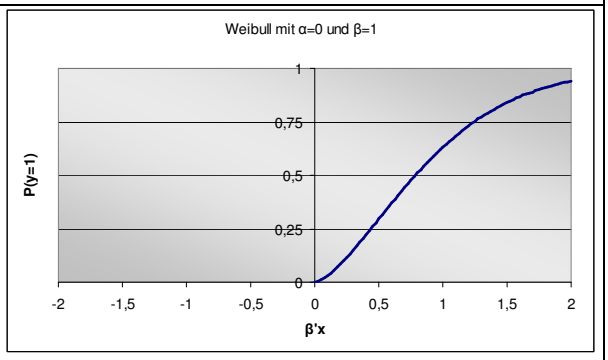


$$P(y_i = 1) = \left(\frac{e^{\mathrm{x_i}\beta}}{1 + e^{\mathrm{x_i}\beta}}\right)^{\gamma} = \left(\Lambda(\mathrm{x_i}\beta)\right)^{\gamma} \quad , \gamma > 0$$

Weibull Dichte- und Verteilungsfunktion



$$f(\mathrm{x}_i\beta) = ab\left(\mathrm{x}_i\beta\right)^{b-1} e^{-a(\mathrm{x}_i\beta)^b}$$



$$P(y_i = 1) = 1 - e^{\left(\frac{-\mathrm{x}_i\beta}{a}\right)^b}$$

2.5 Zusammenfassung und Vergleich LP, LOGIT und PROBIT

Diskrete Entscheidungsmodelle

Allgemeiner Ansatz

$y_i^* = x_i\beta + \varepsilon_i \qquad y_i^* = latent \qquad \varepsilon_i$ ist Zufallsterm $(\mathrm{E}[\varepsilon_i] = 0)$

$y_i = E[y_i] + \varepsilon_i = P[y_i = 1] + \varepsilon_i \qquad y_i = 0,1$

$y_i = F(x_i\beta) + \varepsilon_i \qquad y_i = 0,1$

Lineares Wahrscheinlichkeitsmodell LP

$$y_i = P[y_i = 1] + \varepsilon_i = F(x_i\beta) + \varepsilon_i$$

$$P[y_i = 1] = \begin{cases} 0 & wenn\ x_i\beta \le 0 \\ I(x_i\beta) = x_i\beta & wenn\ 0 < x_i\beta < 1 \\ 1 & wenn\ x_i\beta \ge 1 \end{cases}$$

$P[y_i = 1] = x_i\beta \qquad F^{-1}(P_i) = I^{-1}(P_i) = x_i\beta$

$y_i = x_i\beta + \varepsilon_i$

LOGIT

$$y_i = P[y_i = 1] + \varepsilon_i = F(x_i\beta) + \varepsilon_i$$

$$P[y_i = 1] = \Lambda(x_i\beta) = \frac{1}{1 + e^{-x_i\cdot\beta}} \qquad F^{-1}(P_i) = \Lambda^{-1}(P_i) = \log\frac{P_i}{1 - P_i} = x_i\beta$$

$$logit(P_i) = \Lambda^{-1}(P_i) = \log\frac{P_i}{1 - P_i} = x_i\beta$$

$$y_i = \Lambda(x_i\beta) + \varepsilon_i$$

PROBIT

$$y_i = P[y_i = 1] + \varepsilon_i = F(x_i\beta) + \varepsilon_i$$

$$P[y_i = 1] = \Phi(x_i\beta) = \int_{-\infty}^{x_i\beta} \frac{e^{-t^2/2}}{\sqrt{2\pi}}\,dt \qquad F^{-1}(P_i) = \Phi^{-1}(P_i) = x_i\beta$$

$$probit(P_i) = \Phi^{-1}(P_i) = x_i\beta$$

$$y_i = \Phi(\mathrm{x}_i\beta) + \varepsilon_i$$

2.6 Generalisierte Lineare Modelle (GLM) und binäre Entscheidungsmodelle

Generalisierte lineare Modelle (Verallgemeinerte lineare Modelle) sind eine Verallgemeinerung des klassischen linearen Regressionsmodells. Im klassischen linearen Regressionsmodell (CLR) wird angenommen, dass der stochastische Fehlertreme normalverteilt ist. In generalisierten linearen Modellen kann der Fehlertrem eine Verteilung aus der Klasse der exponentiellen Familie besitzen. Neben der Normalverteilung umfasst diese Klasse auch die Binomial-, Poisson-, Gamma- und inverse Gaußverteilung. Damit werden mit GLM viele Regressionsmodelle in einem einheitlichen methodischen Rahmen formuliert (z.B. Logit-Modelle für binäre Variablen oder CLR mit der Normalverteilungsannahme).

Komponenten Generalisierter Linearer Modelle (GLM)

1. Eine zu erklärende Variable y ist eine Funktion von unabhängigen Kovariablen $x = (x_1, ..., x_K)$. y habe eine Verteilung der exponentiellen Form. Im klasischen linearen Regressionsmodell (CLR) mit $y_i = x_i\beta + \varepsilon_i$ ist ist y normalverteilt, da ε_i als normalverteilt angenommen wird.

2. *Linearer Prädiktor:* Die Verteilung von y wird mit den Kovariablen durch eine lineare Funktion (*linearer Prädiktor*) $\eta_i = x_i\beta$ verbunden.

3. *Responsefunktion:* Der Erwartungswert $E(y) = \mu$ ist eine differenzierbare, monotone und damit invertierbare Funktion des linearen Prädiktors η_i. Der Erwartungswert μ wird über eine *Responsefunktion* $h(\eta) \in [0,1]$ mit dem linearen Prädiktor η_i verknüpft $\mu = h(\eta) = h(x_i\beta)$. Damit können viele Verteilungsfunktionen als Responsefunktion verwendet werden.

4. *Linkfunktion*: Die Verknüpfungsfunktion $g = h^{-1}(\eta) = h^{-1}(x_i\beta)$ ist die Inverse der Responsefunktion mit $\eta = g(x_i\beta)$ und gibt die Stelle $x_i\beta$ an, an der die Verteilungsfunktion $h(x_i\beta)$ einen bestimmten (Wahrscheinlichkeits-) Wert besitzt.

Exponentielle Familie

Die Verteilungsfunktion einer Variablen y gehört zur exponentiellen Familie, wenn sich die Wahrscheinlichkeitsfunktion (pdf) formulieren lässt als

$$f(y, \theta, \psi) = \exp\left(\frac{y\theta - b(\theta)}{a(\psi)} + c(y, \psi)\right)$$

mit

$$E(Y) = \mu = b'(\theta)$$
$$Var(Y) = \sigma^2 = b''(\theta)a(\psi)$$

Beispiele für Verteilungen, die zur exponentiellen Familie gehören

Verteilung $E(Y)=\mu$	θ	ψ	$a(\psi)$	$b(\theta)$	$c(y,\psi)$	$f(y)$
Normalverteilung	μ	σ	ψ^2	$\frac{\theta^2}{2}$	$\frac{-y^2}{2\psi}-\log\left(\sqrt{2\pi\psi}\right)$	$\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)$
Bernoulli-Verteilung	$\log\left(\frac{\mu}{1-\mu}\right)$	–	1	$\log(1+e^{\theta})$	0	$\mu^y(1-\mu)^{1-y}$ mit $y=0$ oder 1
Binomialverteilung	$\log\left(\frac{\mu}{n-\mu}\right)$	–	1	$n\log(1+e^{\theta})$	$\log\binom{n}{y}$	$\binom{n}{y}\left(\frac{\mu}{n}\right)^y\left(1-\frac{\mu}{n}\right)^{n-y}$ mit $y=0,1,...,n$
Poisson-Verteilung	$\log(\mu)$	–	1	$\exp(\theta)$	$-\log(y!)$	$\frac{\mu^y}{y!}\exp(-\mu)$ mit $y=0,1,...$

Quelle: wikipedia

GLM Literatur: Fahrmeir, Kneib und Lang 2009, Kapitel 4

GLM und binäre Entscheidungen

Wie wir gesehen haben, ist der Erwartungswert der 0,1-Variablen y die Wahrscheinlichkeit $E(y_i)=P[y_i=1]=F(x_i\beta)$. Die Verbindung zum GLM ist damit wie folgt gegeben:

Binäres Regressionsmodell: $E(y_i)=P[y_i=1]=F(x_i\beta)$

GLM: $$E(y_i)=\mu=h(\eta)=h(x_i\beta)$$

mit der *Responsefunktion* $h(x_i\beta)=F(x_i\beta)$ und der *Linkfunktion* $g=h^{-1}(x_i\beta)=F^{-1}(x_i\beta)=x_i\beta$.

Für das lineare Wahrscheinlichkeitsmodell, das Logit- und das Probit-Modell gilt dann

LP-Modell

$$P[y_i=1]=F(x_i\beta)=h(x_i\beta)$$
$$=I(x_i\beta)=x_i\beta \qquad F^{-1}(P_i)=h^{-1}(P_i)=g(P_i)=\eta=I^{-1}(P_i)=x_i\beta$$

LOGIT-Modell

$$P[y_i=1]=F(x_i\beta)=h(x_i\beta)$$
$$=\Lambda(x_i\beta)==\frac{1}{1+e^{-x_i\cdot\beta}} \qquad F^{-1}(P_i)=h^{-1}(P_i)=g(P_i)=\eta=\Lambda^{-1}(P_i)=\log\frac{P_i}{1-P_i}=x_i\beta$$

PROBIT-Modell

$$P[y_i=1]=F(x_i\beta)=h(x_i\beta)$$
$$=\Phi(x_i\beta)=\int_{-\infty}^{x_i\beta}\frac{e^{-t^2/2}}{\sqrt{2\pi}}dt \qquad F^{-1}(P_i)=h^{-1}(P_i)=g(P_i)=\eta=\Phi^{-1}(P_i)=z_i=x_i\beta$$

2.7 Umrechnungder geschätzten Koeffizienten zwischen den alternativen Ansätzen LP, LOGIT und PROBIT

Ziel der binären Entscheidungsmodelle ist die Schätzung des Koeffizienten β (Reaktionskoeffizient der erklärenden Variablen x_i)

Die geschätzten Koeffizienten aus den alternativen Ansätzen sind wegen unterschiedlicher Verteilungsannahmen nicht direkt zu vergleichen. Allerdings gibt es die folgenden approximativen Umrechnungen der geschätzten Koeffizienten:

$$Var(LOGIT) = \pi^2 / 3 \qquad \text{Var}\left(\text{PROBIT}\right) = 1$$
$$= 3,29 \qquad \text{standardnormalverteilt}$$

$$L_{\tau(x)} = \frac{\exp(\tau \cdot x)}{1+\exp(\tau \cdot x)} \qquad \text{Beziehung zur Normalverteilung}$$

$$\sigma_L = \sigma_{NV} \qquad \Rightarrow \tau = \pi / 3^{0,5} \qquad \text{aber besser } \tau = 1,6$$
$$\Rightarrow \ \text{L}_{1,6}(x) \approx \Phi(x)$$

$$\beta_L \approx 1,6 \cdot \beta_\Phi$$
$$\beta_\Phi \approx 0,625 \cdot \beta_L$$
$$\beta_{LP} \approx 0,25 \cdot \beta_L \text{ (+0.5 für CONST).}$$

$$\beta_{LP} \approx 0,4 \cdot \beta_\Phi \text{ (+0,5 für CONST).}$$
$$\beta_\Phi \approx 2,5 \cdot \beta_{LP} - 0,5$$
$$\beta_\Phi \approx 2,5 \cdot \beta_{LP} - 1,25 \text{ von CONST (Maddala 1983, S. 23)}$$

2.8 Interpretation der Modellergebnisse – Beispiel Kreditvergabe

Vergleichendes Beispiel: Kreditvergabe in Columbia, South Carolina (Maddala 1983, S. 24)

Untersuchungsziel: Erklärung der Kreditvergabe

Sample: Bewerbungen um Hypothekendarlehen ($i = 1,...,n$), $y^* \triangleq$ Kreditwürdigkeit (latente Variable)

$$y_i^* = x_i \beta + c_i$$
$$y_i = \begin{cases} 1 & \text{Darlehensvergabe } \left(y_i^* > 0\right) \\ 0 & \text{sonst} \end{cases}$$

$$y_i^* = x_{0i}\beta_0 + x_{1i}\beta_1 + x_{2i}\beta_2 + \ldots + x_{ni}\beta_n + \varepsilon_i,$$
$$y_i^* = const + AI \cdot \beta_1 + XMD \cdot \beta_2 + DF \cdot \beta_3 + \ldots + NA \cdot \beta_{10} + \varepsilon_i$$

Ergebnisse der LP, LOGIT und PROBIT Schätzung

Achtung: LP und LOGIT Koeffizienten sind entsprechend oben genannter approximativer Umrechnungsformeln bereits in „PROBIT-Koeffizienten“ transformiert worden.

Variable[a]	Linear prob. model	Logit model	Probit model	x_i^c
AI	1,489 (4,69)[b]	2,254 (4,60)	2,030 (4,73)	0,5
XMD	−1,509 (5,74)	−1,170 (5,57)	−1,773 (5,67)	3,0
DF	0,140 (0,78)	0,563 (0,87)	0,206 (0,95)	1
DR	−0,266 (1,84)	−0,240 (1,60)	−0,279 (1,66)	1
DS	−0,238 (1,75)	−0,222 (1,51)	−0,274 (1,70)	1
DA	−1,426 (3,52)	−1,463 (3,34)	−1,570 (3,29)	0,2
NNWP	−1,762 (0,74)	−2,028 (0,80)	−2,360 (0,85)	0,01
NMFI	0,150 (0,23)	0,149 (0,20)	0,194 (0,25)	0,6
NA	−0,393 (1,34)	−0,386 (1,25)	−0,425 (1,26)	0,5
Const.	0,501	0,363	0,488	1

Note: Total number of observation = 750; number of applications rejected = 250; number of applications accepted = 500. To make the coefficients comparable to one another, we have multiplied the logit coefficients by 0,625 and the coefficients of the linear probability model by 2,5 and then subtracted 1,25 from the constant term, as explained in the text.

[a] *AI, applicant's + coapplicant's incomes (in 10^5 dollars); XMD, debt minus mortgage payment (in 10^3 dollars); DF, 1 if female; DR, 1 if nonwhite; DS, 1 if single; DA, age of house (in 10^2 years); NNWP, (neighborhood % nonwhite) x 10^3; NMFI, neighborhood mean family income (in 10^5 dollars); NA, neighborhood average age of homes (in 10^2 years).*

[b] *t-ratios in parentheses.*

[c] *Example of an individual's x_i values (Ms. Smith).*

Quelle: Maddala 1983, S. 24

Für die folgende Interpretation der Modellergebnisse generell und an diesem Beispiel wird das Ergebnis für Ms. Smith mit den Charakteristika der letzten Spalte herangezogen

a) Wahrscheinlichkeiten

Mrs. Smith möchte ein Darlehen. Wie groß ist die Wahrscheinlichkeit der Darlehensakzeptanz?

		Individuum i (Mrs. Smith)	
AI	Applicant's + coapplicant's incomes (10^5 $)	0,5	(50000 $)
XMD	Debt. mortgage payment (10^3 $)	3,0	(3000 $)
DF	Dummy 1 = female 0 = male	1	
DR	Dummy 1 = nonwhite 0 = white	1	
DS	Dummy 1 = single 0 = not	1	
DA	Age of house (10^2 years)	0,2	(20 Jahre)
NNWP	(neighborhood % nonwhite) x 10^3	0,01	(10 %)
NMFI	neighborhood mean family income/10^5 $)	0,6	(60000 $)
NA	neighborhood average age of homes (10^2years)	0,5	(50 Jahre)
Const.	Constant = 1	1	

Rückrechnung der von Maddala schon umgerechneten Koeffizienten

LP-Modell				**Logit-Modell**			**PROBIT-Modell**	
	ß Beispiel =β_{LP} transformiert auf Probit	**ß original**	**Mrs. Smith (x_i)**	**x_iß**	**ß Beispiel =$\beta_L$transformiert auf Probit**	**ß original**	**x_iß**	**Probit**	**x_iß**
AI	1,489	0,5956	0,5	0,2978	2,254	3,6064	1,8032	2,03	1,015
XMD	-1,509	-0,604	3	-1,811	-1,17	-1,872	-5,616	-1,773	-5,319
DF	0,14	0,056	1	0,056	0,563	0,9008	0,9008	0,206	0,206
DR	**-0,266**	**-0,106**	**1**	**-0,106**	**-0,24**	**-0,384**	**-0,384**	**-0,279**	**-0,279**
DS	-0,238	-0,095	1	-0,095	-0,222	-0,355	-0,355	-0,274	-0,274
DA	-1,426	-0,57	0,2	-0,114	-1,463	-2,341	-0,468	-1,57	-0,314
NNWP	-1,762	-0,705	0,01	-0,007	-2,028	-3,245	-0,032	-2,36	-0,024
NMFI	0,15	0,06	0,6	0,036	0,149	0,2384	0,143	0,194	0,1164
NA	-0,393	-0,157	0,5	-0,079	-0,386	-0,618	-0,309	-0,425	-0,213
Const.	0,501	1,4504	1	1,4504	0,363	0,5808	0,5808	0,488	0,488
Ms Smith nonwhite (xß)				-0,372			-3,737		-4,597
Ms. Smith white (xß)				-0,266			-3,353		-4,318
Prob (Kredit; Ms. Smith nonwhite)				0			0,02327108		2,1453E-06
Prob (Kredit; Ms. Smith white)				0			0,03379713		7,87824E-06
Delta Prob (Kredit; Ms. Smith nonwhite-white)				**-0,106**			**-0,01052605**		**-5,7329E-06**

Rückrechnungen:

$$\beta_{k,L}^{Beispiel} = 0,625\beta_{K,L} \qquad \Rightarrow \beta_{k,L} = 1,6\beta_{k,L}^{Beispiel}$$

$$\beta_{k,LP}^{Beispiel} = 2,5\beta_{k,LP} - 1,25 \; von \; Const. \qquad \Rightarrow \beta_{k,LP} = 0,4\beta_{k,LP}^{Beispiel} + 1,25 \; von \; Const.$$

LP: $\sum_K \beta_K x_K = -0,372$

LOGIT: $\sum_K \beta_K x_K = -3,737$

PROBIT: $\sum_K \beta_K x_K = -4,597$

LP-Modell

$$P[y_i = 1] = \begin{cases} 0 & wenn \; x_i\beta \le 0 \\ x_i\beta & wenn \; 0 < x_i\beta < 1 \\ 1 & wenn \; x_i\beta \ge 1 \end{cases}$$

$$x_i\beta = -0,372 \quad \rightarrow \quad P[y_i = 1] = '-0,372' \le 0$$

$$P[y_i = 1] = 0$$

Ms. Smith nonwhite: keine Chance für einen Kredit

LOGIT-Modell

$$P[y_i = 1] = \frac{1}{1 + \exp(-x_i\beta)}$$

$$x_i\beta = -3,737 \Rightarrow P[y_i = 1] = \left[1 + \exp(3,737)\right]^{-1} = 0,02327$$

Ms. Smith nonwhite: Kreditchance 2 %

PROBIT-Modell

$$P[y_i = 1] = \int_{-\infty}^{x_i\cdot\beta} (2\cdot\pi)^{-1/2} \cdot \exp(-t^2/2)\,dt = \Phi(x_i\beta)$$

$$= \Phi(-4,597) = 2,1453\,10^{-6} = 0,00000214 = 0,000214\%$$

ET: PROBD → Home → NORMAL → Parameter → C → $P = 0,000\ldots$

Ms. Smith nonwhite: Kreditchance: so gut wie keine

b) Marginale Wahrscheinlichkeiten allgemein

Wegen der Nichtlinearität geben die Koeffizienten der binären Entscheidungsmodelle – außer bei dem linearen Wahrscheinlichkeitsmodell – nicht unmittelbar auch die marginalen Effekte – also die Änderung der Wahrscheinlichkeit, wenn sich ein x um eine Einheit verändert - an.

Allgemein werden die marginalen Effekte (Kettenregel) bestimmt durch

$$\frac{\partial E(y \mid x)}{\partial x} = \frac{\partial F(x_i\beta)}{\partial x_i} = \frac{\partial F(x_i\beta)}{\partial (x_i\beta)} \frac{\partial (x_i\beta)}{\partial x_i} = \left(\frac{\partial F(x_i\beta)}{\partial (x_i\beta)}\right)\beta = f(x_i\beta)\beta$$

wobei f(.) die Wahrscheinlichkeitsdichtefunktion korrespondierend zur kumulativen Verteilungsfunktion F(.) ist.

Damit sind die marginalen Effekte abhängig vom Niveau der x-Werte. Oft werden die Mittelwerte der Regressoren (x) oder geeignete Ausgangswerte genommen. Für die Berechnung ist festzuhalten, dass der Skalierungsfaktor $f(x_i\beta)$ der gleiche ist für alle Regressionskoeffizienten β.

Für unsere drei binären Modelle gilt dann allgemein:

LP-Modell

$$\frac{\partial E(y \mid x)}{\partial x} = \left(\frac{\partial (x_i\beta)}{\partial (x_i\beta)}\right)\beta = 1\beta = \beta$$

LOGIT-Modell

$$\frac{\partial E(y \mid x)}{\partial x} = \left(\frac{d\Lambda(x_i\beta)}{d(x_i\beta)} \right) \beta = \lambda(x_i\beta)\beta = \left(\exp(x_i\beta)/(1+\exp(x_i\beta))^2 \right) \beta$$
$$= \Lambda(x_i\beta)\left(1-\Lambda(x_i\beta)\right)\beta$$

PROBIT-Modell

$$\frac{\partial E(y \mid x)}{\partial x} = \left(\frac{\partial \Phi(x_i\beta)}{\partial(x_i\beta)} \right) \beta = \phi(x_i\beta)\beta$$

Beispiel Kreditvergabe (marginale Effekt von Mrs. Smith white auf Mrs. Smith nonwhite):

Wie ändert sich die Wahrscheinlichkeit für eine Darlehensakzeptanz, wenn Mrs. Smith statt white nonwhite wäre?

LP-Modell

$$\frac{\partial E(y \mid x)}{\partial x} = \beta$$
$$= -0,106 = -10,6\%$$

Reduktion der Wahrscheinlichkeit einer Darlehenszusage um 11% falls Mrs. Smith nonwhite statt white wäre.

LOGIT-Modell

Ausgangssituation: Mrs. Smith, white, $x_i\beta = -3,353$

$$\frac{\partial E(y \mid x)}{\partial x_{DR}} = \left(\exp(x_i\beta)/(1+\exp(x_i\beta))^2 \right) \beta$$
$$= 0,032655\beta_{DR} = 0,032655 * (-0,384) = -0,0125 = -1,25\%$$

Reduktion der Wahrscheinlichkeit einer Darlehenszusage um 1% falls Mrs. Smith nonwhite statt white wäre.

PROBIT-Modell

Ausgangssituation: Mrs. Smith, white, $x_i\beta = -4,318$

$$\frac{\partial E(y \mid x)}{\partial x} = \phi(x_i\beta)\beta = \phi(x_{white}\beta)\beta_{DR}$$
$$= \phi(-4,318)(-0,279) = -3,56593 \cdot 10^{-5} \cdot (-0,279)$$
$$= -0,0000099488... = -0,0009948\%$$

Reduktion um ca. 0,00099%.
Aber Achtung: Dummy-Änderung ist keine infinitesimale Änderung! Siehe nächsten Abschnitt.

c) Marginale Wahrscheinlichkeiten bei Dummies – Vergleich zweier Wahrscheinlichkeiten insgesamt

Die obige Berechnung der marginalen Wahrscheinlichkeiten gilt streng genommen nur für stetige Variablen x. Die Auswirkungen von Änderungen der x-Variablen auf die Wahrscheinlichkeit kann generell – und insbesondere für Dummies (nichtstetige Variablen) – über die Differenz der entsprechenden Wahrscheinlichkeiten insgesamt vor und nach Änderung angegeben werden.

Marginale Effekte von Dummy-Variablen

$$\text{Marginaler Effekt} = \mathrm{P}(y = 1 \mid x_{dummy} = 1) - \mathrm{P}(y = 1 \mid x_{dummy} = 0)$$

Beispiel Kreditvergabe (marg. Effekt von Mrs. Smith white auf Mrs. Smith nonwhite):

Wie ändert sich die Wahrscheinlichkeit (insgesamt) für eine Darlehensakzeptanz, wenn Mrs. Smith statt white nonwhite wäre (Delta-Methode)?

$$\text{Marginaler Effekt} = \mathrm{P}(y = 1 \mid nonwhite) - \mathrm{P}(y = 1 \mid white)$$

Lineares Wahrscheinlichkeitsmodell (LP)

Lineare Transformation $P(y = 1) = F(\mathrm{x}_i\beta) = \mathrm{x}_i\beta$.

$$P(y = 1 \mid nonwhite) - P(y = 1 \mid white) =$$
$$x_{nonwhite}\beta - x_{white}\beta = -0,372 + 0,266 = -0,106$$

To be nonwhite‘ reduziert die Wahrscheinlichkeit einer Kreditvergabe um ca. 11 %.

LOGIT-Modell

Nichtlineare Transformation $P(y = 1) = F(x_i\beta) = [1 + \exp(-x_i\beta)]^{-1}$

β_k ist kein direkter Wahrscheinlichkeitsbeitrag, aber wegen der monotonen Transformation des LOGIT-Ansatzes sind über das Vorzeichen jedoch Richtungsaussagen direkt ablesbar.

'To be nonwhite' $DR = 1$

Achtung: Koeffiziententransformation rückgängig machen: $\beta_{k,L} = -0,240 / 0,625 = -0,384$

Ausgangssituation ist Ms. Smith mit ihrem $x_i\beta$ aber DR=0 (white)

$$\text{DR} = 0: \quad PROB^V = \left[1 + \exp(3,353)\right]^{-1} = 0,03379$$

$$\text{DR} = 1: \quad PROB^N = \left[1 + \exp(3,737)\right]^{-1} = 0,02327$$

$$\Delta PROB = PROB^N - PROB^V = -0,01052 = -1,05\%$$

(abhängig von $x_i\beta$ vorher!)

$\Rightarrow \beta_n$ negativ, auch $\Delta\, PROB$ negativ

‚To be nonwhite' reduziert mit dem LOGIT-Modell die Wahrscheinlichkeit einer Kreditvergabe von Ms. Smith um 1,05 % gegenüber ‚To be white'.

PROBIT-Modell

Nichtlineare Transformation

$$P[y = 1] = F(x_i\beta) = \int_{-\infty}^{x_i \cdot \beta} (2\pi)^{-1/2} \cdot \exp(-t^2/2)dt = \Phi(x_i\beta)$$

β_K ist kein direkter Beitrag, aber monotone Transformation $\Rightarrow$ Richtungsaussage ist somit über das Vorzeichen wieder direkt ablesbar.

z. B. 'To be nonwhite' $DR = 1$ $\Rightarrow$

Achtung: Keine Koeffiziententransformation nötig, da das PROBIT-Modell das Bezugsmodell ist.

$$\text{DR} = 0: \quad PROB^V = \Phi(x_{nonwhite}\beta) = \Phi(-4,597) = 2,1453\ 10^{-6}$$

$$\text{DR} = 1: \quad PROB^V = \Phi(x_{white}\beta) = \Phi(-4,318) = 7,8782\ 10^{-6}$$

$$\Delta PROB = PROB^N - PROB^V = -5,7329\ 10^{-6} = -0,000005732... = -0,0005732\%$$

(abhängig von $x_i\beta$ vorher!)

$\Rightarrow \beta_n$ negativ, auch $\Delta\, PROB$ negativ

D.h. ‚To be nonwhite' reduziert mit dem PROBIT-Modell die Wahrscheinlichkeit einer Kreditvergabe von Ms. Smith um -0,00057% gegenüber ‚To be white'.

Insgesamt:

Marginale Wahrscheinlichkeiten **Beispiel Kreditvergabe Mrs. Smith nonwhite statt white**		
	infinitesimal	Delta
LP-Modell	-10,6%	-10,6%
LOGIT-Modell	-1,25%	-1,05%
PROBIT-Modell	-0,000995%	-0,00057%

3 Schätzen von binären Entscheidungsmodellen

Wir betrachten ein Ereignis $j = \begin{cases} 1 & \text{wird gewählt} \\ 0 & \text{wird nicht gewählt} \end{cases}$

Daten für binäre Entscheidungsmodelle können vorliegen als:

- Gruppierte Daten: best. % aller Beobachtungen $\hat{=}$ 1
- Individuelle Daten: jede Beobachtung mit der Information 0,1

In fast allen Anwendungen binärer Entscheidungsmodelle wird nach dem Effekt von x auf die Eintrittswahrscheinlichkeit $P[y_i = 1 | x]$ gesucht. Mit dem latenten Indexfunktions-Ansatz wird der Effekt von x auf y^* , und im binären Regressionsansatz auf y betrachtet. Die *Richtung* des Effektes von x_j in $E(y^* | x) = x_i\beta$ und in $E(y | x) = P[y_i = 1 | x] = F(x_i\beta)$ ist immer gleich (Wooldridge 2009, S. 577). Allerdings hat die latente Variable (z.B. als Nutzendifferenz) keine eindeutige Maßeinheit. Deshalb steht nicht die Größe eines jeden β_j im Vordergrund sondern eher der marginale Effekt von x_j .

3.1 Schätzen mit gruppierten Daten

Aus allen Beobachtungen (n) werden g = 1, ..., G, Gruppen mit jeweils n_g Beobachtungen gebildet.

Gruppenbildung

Zur besseren Schätzung sind möglichst viele Gruppen zu bilden: z.B. Männer, Frauen nach Altersklassen, sozioökonomischen Kategorien etc. bilden.

Für jede Gruppe:

$$\text{Stichprobenanteil } \hat{\pi}_g = \frac{\text{Anzahl der Beobachtungen mit } y_i = 1}{\text{Alle Beobachtungen der Gruppe } (n_g)}$$

Diese Stichprobenanteile $\hat{\pi}_g$ (beobachtete relative Häufigkeiten) gehen in unterschiedlicher Weise als Schätzer für *P* (Auswahlwahrscheinlichkeit) ein. Gesucht: $\hat{\beta}$ als Schätzer für β.

LP-Modell

$$P[y_i = 1] = \begin{cases} 0 & x_i\beta < 0 \\ I(x_i\beta) = x_i\beta & 0 \le x_i\beta \le 1 \\ 1 & x_i\beta > 1 \end{cases}$$

Gruppenschätzer:

$$P[y_i = 1] = F(x_g\beta) = I(x_g\beta) = \hat{\pi}_g = x_g\beta \quad \textit{lineare Schätzung}$$

$$I^{-1}(\hat{\pi}_g) = \hat{\pi}_g^{-1} = \hat{\pi}_g = x_g\beta \qquad OLS / WLS \Rightarrow \hat{\beta}$$

d.h. Schätzung einer Ausgleichsebene aus g = 1,...,G gruppierten Beobachtungen: LHS ('left hand side') : $\hat{\pi}_g$ z.B. mit OLS (wegen Heteroskedastizität besser mit 'weighted least squares' (WLS) $\Rightarrow \hat{\beta}$)

LOGIT-Modell:

$$P[y_i = 1] = F(x_i\beta) = \Lambda(x_i\beta) = \left[1 + \exp(-x_i\beta)\right]^{-1}$$

Nichtlineare LOGIT-Transformation als geschlossene Lösung $F^{-1}(x_i\beta)$:

$$F^{-1}(P_i) = \Lambda^{-1}(P_i) = \log\left(\frac{P(x_i\beta)}{1 - P(x_i\beta)}\right) = x_i\beta$$

Gruppenschätzer:

$$P[y_i = 1] = F(x_g\beta) = \Lambda(x_g\beta) = \hat{\pi}_g = \left[1 + \exp(-x_g\beta)\right]^{-1} \quad \textit{nichtlineare Schätzung oder}$$

$$\Lambda^{-1}(\hat{\pi}_g) = \log\left(\frac{\hat{\pi}_g}{1 - \hat{\pi}_g}\right) = x_g\beta \qquad OLS / WLS \Rightarrow \hat{\beta}$$

Zur Interpretation: werden Dummy-Variablen als erklärende Größen im LOGIT Modell verwendet, dann wird die ausgelassene Referenz mit 0 codiert. Wird sie dagegen mit -1 codiert (‚effect coding'), dann können die verbleibenden Terme als Abweichungen vom Mittelwert (allerdings dem Mittelwert über die Kategorien, nicht über die Individuen) interpretiert werden (Alba 1988).

PROBIT-Modell:

$$P[y_i = 1] = \Phi(x_i\beta) = \int_{-\infty}^{x_i\beta/\sigma} \phi(z)\,dz$$

Frühe Verwendung in der Biometrie: 'bioassay' (Finney 1971, Cox 1970) siehe Greene 1990 S. 666 ff für weitere Erläuterungen.

Gruppenschätzer :

$$P[y_i = 1] = F(x_g\beta) = \Phi(x_g\beta) = \hat{\pi}_g = \int_{-\infty}^{x_i\beta} \frac{e^{-t^2/2}}{\sqrt{2\pi}}\,dt \quad \textit{nichtlineare Schätzung oder}$$

$$\Phi^{-1}(\hat{\pi}_g) = z_g = x_g\beta \qquad OLS\,/\,WLS \Rightarrow \hat{\beta}$$

Aus dem PROBIT-Integralansatz gibt es keine geschlossene Lösung $F^{-1}(P_i) = \Phi^{-1}(P_i) = z$, so dass $F^{-1}(\hat{\pi}_g) = x_g\beta$ direkt gebildet werden könnte.

Zusammenfassung Schätzung mit gruppierten Daten:

Regressionsansatz: $y_{LHS} = x_g\beta + \varepsilon_g$

jeweils zu schätzen mit OLS bzw. GLS, wobei

Allgemein	$\hat{y}_{LHS} = P(x_g\hat{\beta}) = F(x_g\hat{\beta})$
LP-Modell	$\hat{y}_{LHS} = \hat{\pi}_g^{-1} = \hat{\pi}_g$
LOGIT-Modell	$\hat{y}_{LHS} = \log[\hat{\pi}_g\,/(1-\hat{\pi}_g)]$
PROBIT-Modell	$\hat{y}_{LHS} = z_g = \Phi^{-1}(\hat{\pi}_g)$

Beispiele: Schätzung mit gruppierten Daten

a) NOKJA-Handykauf

Für die Untersuchung des Käuferprofils habe die Firma NOKJA eine Umfrage mit n=1000 Befragten durchgeführt. Ein NOKJA-Handykauf ist das beobachtbare Ereignis. Als Gruppierungsmerkmal werden Alters- und Regionalklassen gebildet (z.B. 14 Altersklassen und zwei Regionen =>n=28). Die folgende Tabelle enthält für die ersten beiden Gruppen g=1,2 für das LP-, das LOGIT- und das PROBIT-Modell die entsprechenden linken Seiten (LHS) der Regressionsgleichungen und die gruppenspezifischen Mittelwerte der erklärenden Variablen (rechte Seite, RHS) für eine dann folgende OLS- bzw. GLS-Schätzung aus

$$y_{LHS} = x_g\beta + \varepsilon_g$$

Im gruppierten Ansatz müssen auch die erklärenden Variablen jeweils die Gruppe repräsentieren (z.B. als arithmetisches Mittel von x_i).

Gruppe g	$y_g = 1$	n_g	y_{LHS} (linke Seite) LP	LOGIT	PROBIT	x (rechte Seite) Mittelwerte $\bar{x}_{1g}$ (=age)	$\bar{x}_{2g}$ (=income)
1	45	60	$\hat{\pi}_1 = 0{,}75$	$\log\left(\frac{0{,}75}{1-0{,}75}\right) = 1{,}099$	$z_g = \Phi^{-1}(0{,}75) = 0{,}675$	32	5430,-
2	6	20	$\hat{\pi}_2 = 0{,}3$	$\log\left(\frac{0{,}3}{1-0{,}3}\right) = -0{,}847$	$z_g = \Phi^{-1}(0{,}3) = -0{,}524$	24	3250,-
…							

(z Werte mit Stata: z.B. display invnormal(0.3) resultiert in -0.524)

$$\text{z.B.: } \hat{\pi}_1 = \frac{Anzahl\left[y_{ig} = 1\right]}{n_1} = \frac{45\,(Anzahl\ der\ Handykäufer)}{60\,(Alle\ in\ der\ Gruppe\ 1)} = 0{,}75$$

b) Gruppierte Schätzung: Anwendung bei medizinischen Versuchen

Frühe Anwendung in der Biologie und Medizin:
'Proband' = Versuchstier *i* erhält Dosis x_g

x_i^* sei Verträglichkeitstoleranz

$$\begin{aligned}\text{Prob}\left[y_{ig} = 1\right] &= \text{Prob}\left[x_g < x_g\beta\right] \\ &= \Phi\left(x_g\beta\right) \\ &= \pi_g\end{aligned}$$

Dosis x_g wird nun variiert über die Gruppen → Variation von P_g Anteil der Tiere, die nicht überleben (z.B. sinnloser 'LD50-Test' 'lethal dosis 50': 50 % aller Versuchstiere sterben bei dieser Dosis!).

Ausgehend von der Normalverteilung $P_g \approx \Phi\left(x_g\beta\right) = \int_{-\infty}^{x_g\beta} \phi(z)\,dz$

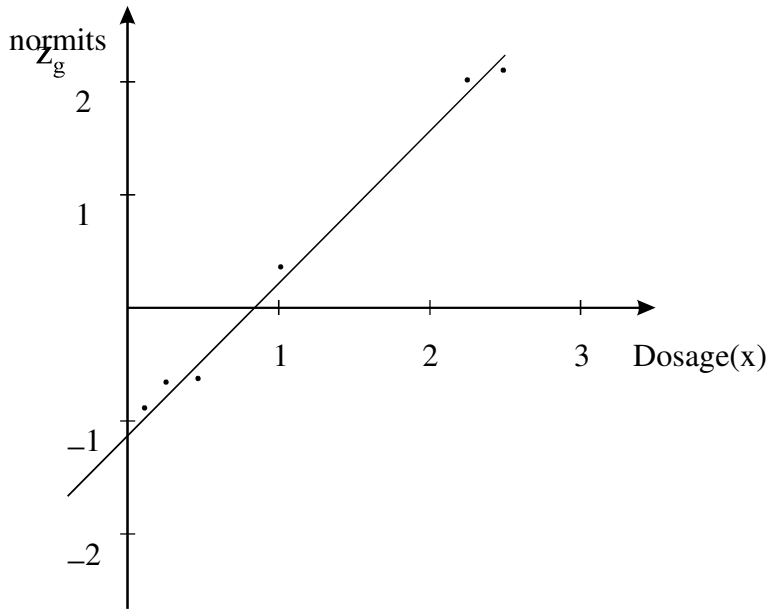
Schätzung von $\hat{\beta}$ mit

$$z_g = \Phi^{-1}\left(\pi_g\right) = x_g\beta \quad mit \quad \pi_g = \Phi\left(x_g\beta\right)$$

$$z_g = '\,normits\,'$$

$$\Phi^{-1}\left(\pi_g\right) = \textit{Normalverteilungstabellen 'backward', oder mit ET, Stata etc.}$$

Unterschiedliche Dosierungen ergeben unterschiedliche Anteile $\hat{P}_g = \hat{\pi}_g$ mit den z_g (*normits*) für die Schätzung von β.



Ausgleichsebene (z.B. durch OLS) durch die Punktwolke (x_g, z_g)

(Normalverteilter Fehler $\hat{\pi}_g = \pi_g + \varepsilon_g$) $\Rightarrow \hat{\beta}$

Insgesamt: Gruppierte Schätzungen von β sind nur approximativ, verzerrt; kurz: problematisch und abhängig von der Gruppenbildung. Besser: Schätzung über den Individualansatz.

3.2 Schätzen mit Individualdaten

Nun sind nicht mehr Anteilswerte $\hat{\pi}_i$ der Ereigniseintritte, sondern für jede Beobachtung der Ereigniseintritt y_i gegeben bzw. beobachtet (i = 1,...,n)

$$y_i = \begin{cases} 1 & \text{falls } y_i^* = x_i\beta + \varepsilon_i > a(=0) \quad \textit{Anspruchsniveau} \\ 0 & \text{sonst} \end{cases}$$

Wie ist der Parametervektor β der Auswahlwahrscheinlichkeit $P[y_i = 1] = F(x_i\beta)$ zu schätzen?

Allgemeiner Ansatz

$y_i^* = x_i\beta + \varepsilon_i$ $\quad y_i^* = \textit{latent}$ $\quad \varepsilon_i$ ist Zufallsterm $(E[\varepsilon_i] = 0)$

$y_i = E[y_i] + \varepsilon_i = P[y_i = 1] + \varepsilon_i$ $\quad y_i = 0,1$

$y_i = F(x_i\beta) + \varepsilon_i$ $\quad y_i = 0,1$

$P_i = F(x_i\beta)$

Lineares Wahrscheinlichkeitsmodell LP

$$P[y_i = 1] = \begin{cases} 0 & wenn\ x_i\beta \le 0 \\ I(x_i\beta & wenn\ 0 < x_i\beta < 1 \\ 1 & wenn\ x_i\beta \ge 1 \end{cases}$$

$$P[y_i = 1] = I(x_i\beta) = x_i\beta$$

LOGIT

$$y_i = P[y_i = 1] + \varepsilon_i = F(x_i\beta) + \varepsilon_i$$

$$P[y_i = 1] = \Lambda(x_i\beta) \qquad F^{-1}(P_i) = \Lambda^{-1}(P_i) = \log\frac{P_i}{1-P_i} = x_i\beta$$

$$P[y_i = 1] = \Lambda(x_i\beta)$$

PROBIT

$$y_i = P[y_i = 1] + \varepsilon_i = F(x_i\beta) + \varepsilon_i$$

$$P[y_i = 1] = \Phi(x_i\beta) \qquad F^{-1}(x_i\beta) = \Phi^{-1}(P_i) = x_i\beta$$

$$P[y_i = 1] = \Phi(x_i\beta)$$

GOMPIT

$$y_i = P[y_i = 1] + \varepsilon_i = F(x_i\beta) + \varepsilon_i$$

$$P[y_i = 1] = F(x_i\beta) = G(x_i\beta) = e^{-e^{x_i\beta}} \qquad F^{-1}(P_i) = G^{-1}(P_i) = \ln\ln(1/P_i) = x_i\beta$$

$$P[y_i = 1] = G(x_i\beta)$$

Der Parametervektor β der Auswahlwahrscheinlichkeit $P[y_i = 1] = F(x_i\beta)$ kann nun generell mit verschiedenen Methoden geschätzt werden:

- **Maximum score** – Schätzer (Manski 1975)

  $max_\beta\ S_{N\alpha}(\beta) = \frac{1}{n}\sum_i \left[z_i - (1-2\alpha)\, sgn(x_i\beta) \right]$

  α = *vorbestimmtes Quantil*

  $z_i = 2y_i - 1 \qquad (z_i = -1 \quad wenn \quad y_i = 0)$

  Nonparametric estimation
- **Minimum Chi-Quadrat –** Schätzer (Maddala 1983, S. 39)

  $GLS \quad für \quad \hat{p}_i = x_i\beta + \varepsilon_i$

  mit $w_i = \left\{ n_i \left[\hat{p}_i (1 - \hat{p}_i) \right] \right\}$

  Diskriminanz-Analyse (Maddala 1983, S. 27)
- **Maximum Likelihood** (ML) – Schätzer (Greene 2000, S. 123 ff)

Der am meisten verwendete Ansatz ist der Maximum Likelihood (ML)-Ansatz:

Maximum Likelihood (ML)-Schätzer

Prinzip: Es wird der Parameter θ einer Grundgesamtheits-Verteilung gesucht, der die beobachtete Stichprobe mit größter Wahrscheinlichkeit liefert. Wird nun jede Beobachtung als eine einzelne Ziehung aus einer Bernoulli Verteilung (binomial mit einem Zug) betrachtet, dann wird mit der Erfolgswahrscheinlichkeit $F(x_i\beta)$ und unabhängigen Beobachtungen dann die gemeinsame Wahrscheinlichkeit, die Likelihood Funktion formuliert mit:

Likelihood-Funktion:

$$\begin{aligned} L(\beta,\theta) &= \prod_i P(y_i \mid x_i) \\ &= \prod_{y_i=1} F(x_i\beta) \prod_{y_i=0} \left[1-F(x_i\beta)\right] \\ &= \prod_i \left[F(x_i\beta)\right]^{y_i} \left[1-F(x_i\beta)\right]^{1-y_i} \\ \ln L &= \sum_i \left[y_i \ln F(x_i\beta) + (1-y_i) \ln\left(1-F(x_i\beta)\right)\right] \end{aligned}$$

Bedingung 1. Ordnung für max L(.) = max ln L(.)

$$\frac{\partial \ln L}{\partial \beta} = \sum_i \left[\frac{y_i f_i}{F_i} + (1-y_i)\frac{-f_i}{(1-F_i)} \right] x_i = 0$$

wobei $f_i = f(x_i\beta)$ *die pdf*
$F_i = F(x_i\beta)$ *die cdf*

Außer bei dem linearen Wahrscheinlichkeitsmodell sind die Bedingung 1. Ordnung-Gleichungen nichtlinear und benötigen eine iterative Lösung, die dann den Schätzer für β hervorbringt.

Amemiya (1985, S. 273-274) hat gezeigt, dass unter praktischen Bedingungen die Likelihood Funktion global konkav ist und damit die Eindeutigkeit der ML-Schätzungen sichert. Die Schätzungen sind konsistent (=asymptotisch erwartungstreu), asymptotisch normal verteilt und asymptotisch effizient. Weil die ML-Schätzung auf der Verteilung von y gegeben x basiert, wird die Heteroskedastizität von $\mathrm{Var}(y|x)$ automatisch berücksichtigt (Wooldridge 2009, S. 578).

Generell liefert ein ML-Schätzer also konsistente und asymptotisch effiziente Schätzergebnisse, die dann zum Testen von Hypothesen normal verteilter Koeffizientenschätzer verwendet werden können.

Maximum Likelihood Ansätze für LP, LOGIT und PROBIT:

Lineares Wahrscheinlichkeitsmodell

$$F(x_i\beta) = x_i\beta$$

Likelihood:

$$L(\beta) = \prod_i (x_i\beta)^{y_i} \left[1-(x_i\beta)\right]^{1-y_i}$$

bzw.

$$\ln \quad L = \sum_i \left\{ y_i \ln(x_i\beta) + (1-y_i) \ln\left[1-(x_i\beta)\right]\right\}$$

$$\ln \quad L = \sum_{y_i=1} y_i \ln(x_i\beta) + \sum_{y_i=0} \ln\left[1-(x_i\beta)\right]$$

Bedingung 1. Ordnung:

$$\frac{\partial \ln L}{\partial \beta} = \sum_{y_i=1} \frac{y_i}{x_i\beta} x_i - \sum_{y_i=0} \frac{1}{(1-x_i\beta)} x_i = 0$$

β als Lösung eines linearen Gleichungssystems (OLS): $\beta = (X'X)^{-1}X'y$.

LOGIT

$$F(x_i\beta) = \Lambda(x_i\beta) = \left[1+\exp(-x_i\beta)\right]^{-1}$$

Likelihood:

$$L(\beta) = \Pi_i \left[\Lambda(x_i\beta)\right]^{y_i} \left[1-\Lambda(x_i\beta)\right]^{1-y_i}$$

bzw.

$\ln L = \sum_i y_i (x_i\beta) - \sum_i \ln\left(1+e^{x_i\cdot\beta}\right)$ nach Vereinfachung

Bedingung 1. Ordnung

$$\frac{\partial \ln L}{\partial \beta} = \sum_i \left[y_i - \Lambda(x_i\beta)\right] x_i = 0$$

β als Lösung eines nichtlinearen Gleichungssystems

PROBIT

$$F(x_i\beta) = \Phi(x_i\beta) = \int_{-\infty}^{x_i\cdot\beta/\sigma} (2\pi)^{-\frac{1}{2}} e^{-\frac{t^2}{2}} dt$$

Likelihood

$$L\left(\beta, \theta = \sigma^2\right) = \prod_i \left[\Phi(x_i\beta)\right]^{y_i} \left[1-\Phi(x_i\beta)\right]^{1-y_i}$$

bzw.

$$\ln L = \sum_{y_i=1} \ln \Phi_i + \sum_{y_i=0} \ln(1-\Phi_i)$$

Bedingung 1. Ordnung

$$\frac{\partial \ln L}{\partial \beta} = \sum_{y_i=1} \frac{\varphi_i}{\Phi_i} x_i + \sum_{y_i=0} \frac{-\varphi_i}{\Phi_i} x_i = 0$$

mit Dichtefunktion

$$\varphi_i = (2\pi)^{-1/2} \cdot e^{-(x_i\beta)^2/2} = \frac{d\Phi(x_i\beta)}{d(x_i\beta)}$$

β als Lösung eines nichtlinearen Gleichungssystems (Achtung: Lösung $\beta = \beta / \sigma^2$)

Für die geschätzten β aus LP, LOGIT oder PROBIT gilt:

Kovarianzmatrix von $\hat{\beta}$

über die Informationsmatrix

$$I(\hat{\beta}) = \left(\frac{\partial^2 \ln L}{\partial \beta \, \partial \beta'} \right)$$

wobei $\frac{\partial^2 \ln L}{\partial \beta \partial \beta'}$ = Hesse'sche Matrix der 2. Ableitungen ist, wird die Kovarianzmatrix von $\hat{\beta}$ gebildet aus

$$Cov(\hat{\beta}) = \left(-E \frac{\partial^2 \ln L}{\partial \beta \, \partial \beta'} \right)^{-1}.$$

Die asymptotische Varianz-Kovarianzmatrix von $\hat{\beta}$ des ML-Schätzers kann also mit der Hesse'schen Matrix der zweiten Ableitungen bei $\hat{\beta}$ berechnet werden.

Numerische Lösung der nichtlinearen Gleichungssysteme des LOGIT- und PROBIT-Ansatzes

Wie deutlich wird, sind sowohl bei PROBIT als auch bei LOGIT die nach $\hat{\beta}$ zu lösenden Gleichungssysteme nichtlinear:

LOGIT: geschlossen, keine Integrale $\left(\Lambda_i = \left[1 + \exp(-x_i \cdot \beta) \right]^{-1} \right)$

PROBIT: nicht geschlossen, Integrale jeweils zu lösen $\Rightarrow$ rechenintensiv

Außer für das lineare Wahscheinlichkeitsmodell ist das Gleichungssystem des ML-Ansatzes im allgemeinen nicht linear. Es gibt keine geschlossene, analytische Lösung.

Notwendig ist eine iterative Lösung: z.B. Newton-Raphson, Method of scoring,…

Newton-Raphson Methode für die Lösung nichtlinearer Gleichungssysteme (vgl. 4.3)

geg. Startwert β^0 z.B. aus LP-Modell $y_i = x_i\beta$ oder $\pi_i = x_i\beta$

$$\beta^1 = \beta^0 - \lambda\left(\frac{\partial^2 \ln L}{\partial\beta\,\partial\beta'}\right)^{-1}\frac{\partial \ln L}{\partial\beta}$$

wobei λ = Schrittweite

$$\frac{\partial^2 \ln L}{\partial\beta\,\partial\beta'} = \text{Hesse'sche Matrix der 2. Ableitungen}$$

Andere iterative Lösungsmethoden als Newton-Raphson wie 'Quadratic Hill-Climbing' oder 'Davidon Fletcher Powell' benötigen noch die 2. Ableitungen an jeder Stelle (Iteration) $x_i\beta$:

LOGIT: $$\frac{\partial^2 \ln L}{\partial\beta\,\partial\beta'} = -\sum_i x_i \cdot \Lambda_i\left(1-\Lambda_i\right)\cdot x_i'$$

PROBIT: $$\frac{\partial^2 \ln L}{\partial\beta\,\partial\beta'} = \sum_i \frac{\left[\phi\left(x_i\beta\right)\right]^2}{\Phi\left(x_i\beta\right)\left[1-\Phi\left(x_i\beta\right)\right]} x_i x_i'$$

Im Allgemeinen: Hesse'sche Matrix (PROBIT, LOGIT) = negativ definit[1] $\Rightarrow$ eindeutige Lösung (Konvergenz). Da in manchen Fällen die Hesse'sche Matrix der 2. Ableitungen (die die Suchrichtung angibt) etwas schwierig zu berechnen ist, wird die Hesse'sche Matrix im Newton-Raphson Verfahren ersetzt:

- Method of Scoring: H ersetzt durch E(H) = Information Matrix
- BHHH Method (Berndt, Hall, Hall and Hausman (1974)): Als Approximation der Information Matrix H: $\sum_{i=1}^{n}\frac{\partial \ln L}{\partial x}\frac{\partial \ln L'}{\partial x}$.

'Discrete choice' und Newton-Raphson:

$$\log L(\beta) = \sum_i f\left(x_i, y_i, \beta\right)$$

Notwendige Bedingung $$\frac{\partial \log L}{\partial\beta} = 0 = g\left(\beta\right)$$ Gradient

lineare Taylor-Entwicklung um β^0 ergibt Newton-Raphson:

$$g\left(\beta\right) = g\left(\beta^0\right) + H\left(\beta^0\right)\left(\beta-\beta^0\right)$$

$$\Rightarrow \beta \approx \beta^0 - \left[H\left(\beta^0\right)\right]^{-1} g\left(\beta^0\right)$$

LOGIT:
$$g_i = \left(y_i - \text{Prob}\left[y_i = 1\right]\right)x_i$$
$$H_i = -P_i\left(1-P_i\right)x_i x_i'$$

[1] Eine quadratische Matrix A ist negativ definit, wenn die quadratische Form x'Ax<0 für alle x ungleich Null ist.

PROBIT

$$g_i = \left[y_i \left(\phi / \Phi \right) + \left(1 - y_i \right) \left(-\phi / \left(1 - \Phi \right) \right) \right] x_i = \gamma_i x_i$$
$$H_i = -\gamma_i \left(\beta' x_i + \gamma_i \right) x_i x_i'$$

POISSON

$$g_i = \left(y_i - E\left(y_i \right) \right) x_i$$
$$H_i = \exp\left(\beta' x_i \right) x_i x_i'$$

3.3 Exkurs: Newton-Raphson-Methode zur Lösung nichtlinearer Optimierungsprobleme

Problem:

Maximum einer Funktion f(x) (z.B. Maximum Likelihood) → Lösung über die Nullstellen der 1. Ableitung

Aufgabe:

Suche Nullstelle einer Funktion

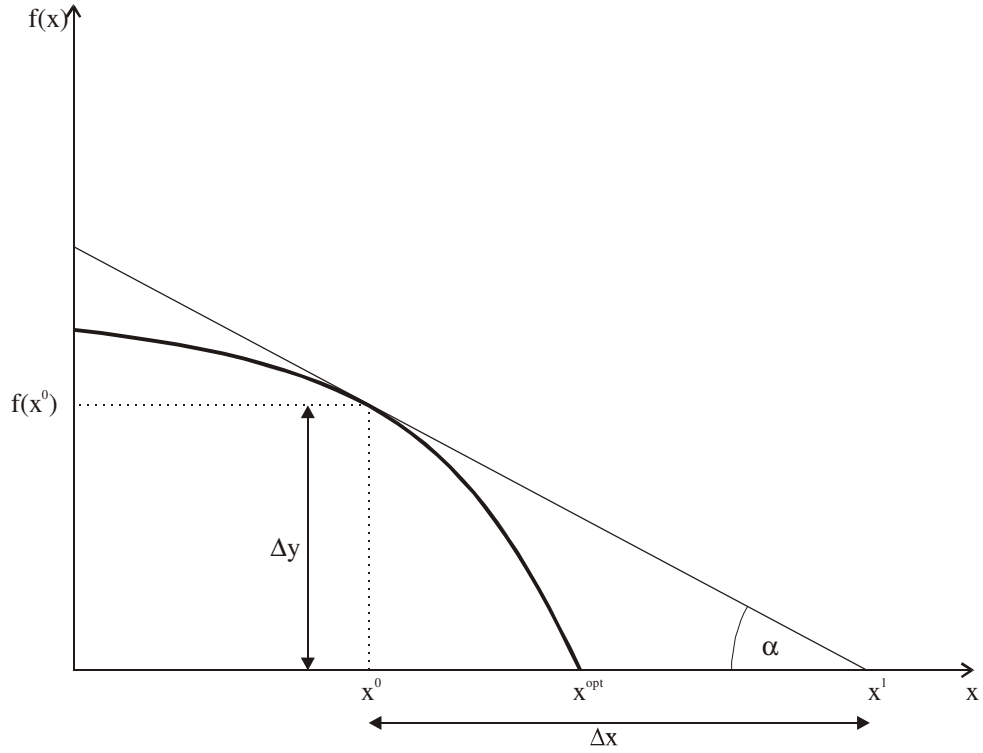
- allgemein: $f(x) = 0 \Rightarrow x^{null}$
- bei Maximierungsproblem: $f'(x) = 0 \Rightarrow x^{opt}$ *mit* $f\left(x^{opt} \right) = \max$ der Ausgangsfunktion

Lösung:

Iteratives Verfahren nach Newton-Raphson zur allgemeinen Nullstellenbestimmung

$$\beta^1 = \beta^0 - \lambda \left(\frac{\partial^2 \ln L}{\partial \beta \, \partial \beta'} \right)^{-1} \frac{\partial \ln L}{\partial \beta}$$

Ableitung des allgemeinen Newton-Ansatzes:



$$\frac{\Delta y}{\Delta x} = \tan\alpha = f'\left(x^0\right)$$

$$\frac{f\left(x^0\right)}{x^0 - x^1} = f'\left(x^0\right) \qquad \rightarrow \quad \frac{f\left(x^0\right)}{f'\left(x^0\right)} = x^0 - x^1$$

$$x^1 = x^0 - \frac{1}{f'\left(x^0\right)} f\left(x^0\right)$$

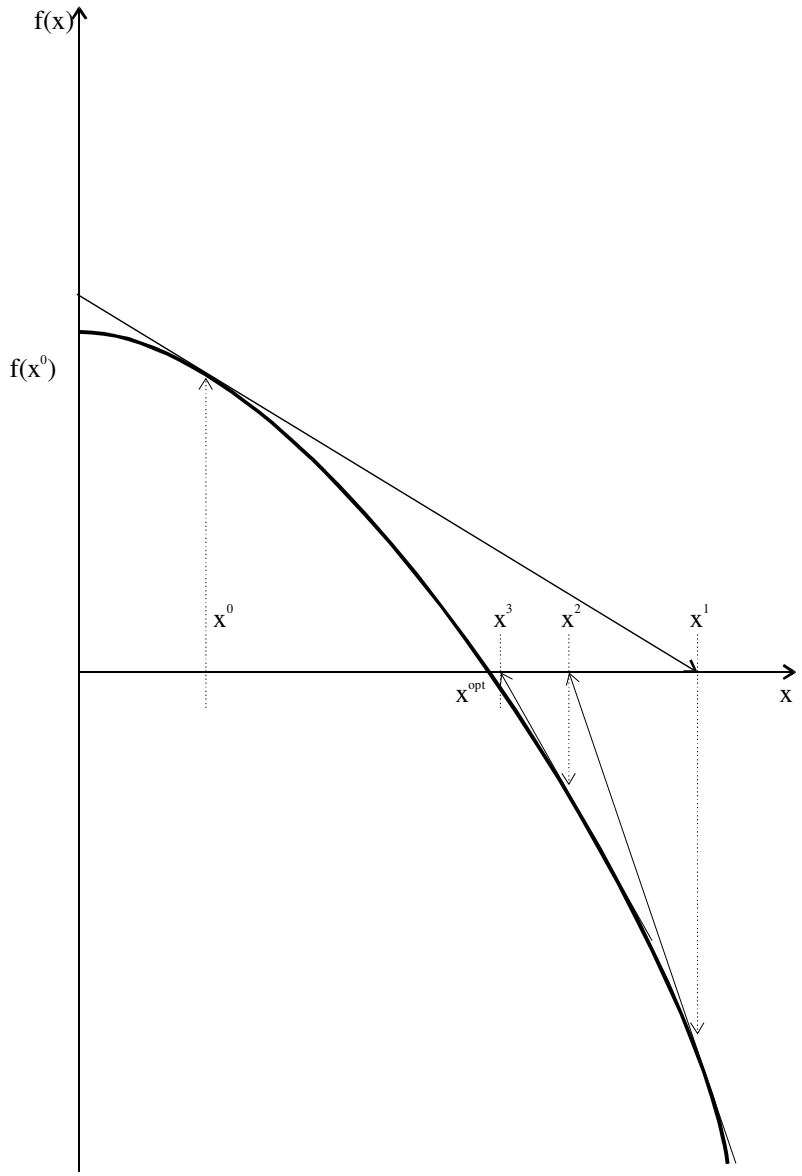
Allgemeiner Ansatz Max/Min einer Funktion:
$\Rightarrow$ Suche Nullstelle der 1. Ableitungen

$$x^1 = x^0 - \frac{1}{f''\left(x^0\right)} f'\left(x^0\right) \quad bzw.$$

$$x^1 = x^0 - \left(\frac{\partial^2 f(x)}{\partial x^2}\right)^{-1}_{x^o} \cdot \frac{\partial f(x)}{\partial x} = x^0 - H^{-1}_{x^0} \cdot g_{x^o}$$

$$g_x = \frac{\partial f(x)}{\partial x} = 'score'$$

$H^{-1}_{x^0}$ = Hesse'sche Matrix der 2. Ableitungen = 'Information'



Schnitt der Tangente von x^0 mit der x-Achse ergibt x^1 etc., Schnitt der Tangenten $f(x^1)$ mit der x-Achse ergibt x^2... Nullstelle $f(x^0)=0$ gefunden.

Beispiel: Newton-Raphson

a) Eindimensional

Maximum gesucht von $f(x)=-x^3+x^2$

=> Nullstelle(n) gesucht der ersten Ableitung $f'(x)$

$$f(x) = -x^3 + x^2$$
$$f'(x) = -3x^2 + 2x$$
$$f''(x) = -6x + 2$$
$$x^1 = x^0 - \frac{1}{f''(x^0)} \cdot f'(x^0)$$
$$x^1 = x^0 - \frac{-3x^2 + 2x}{-6x + 2}$$

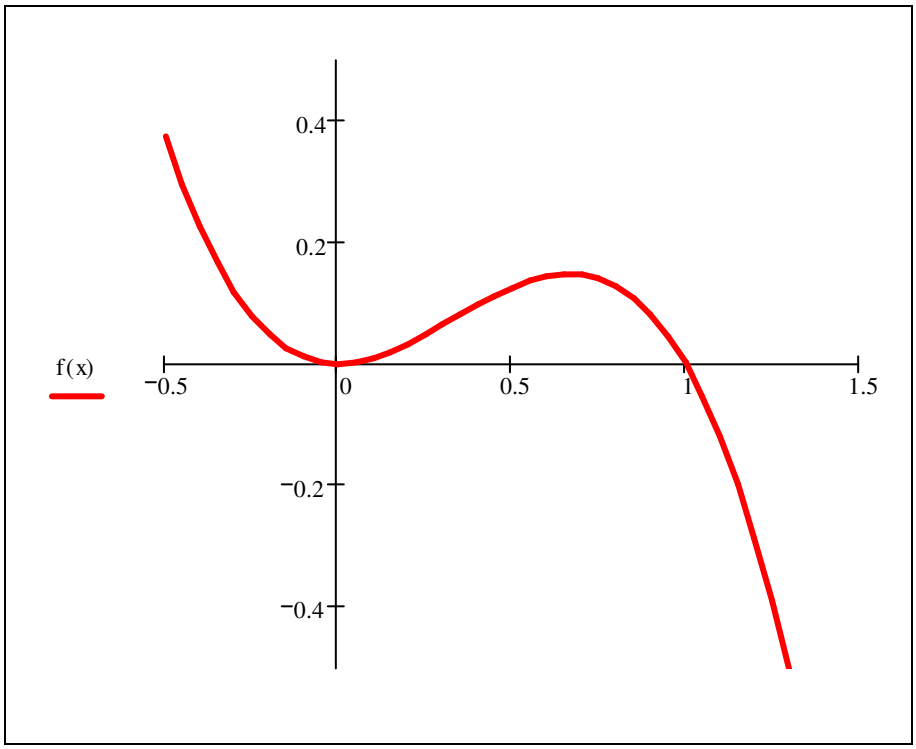
Iteration	x	$f(x)$	$f'(x)$	$f''(x)$
0	4,000	–48,0000	–40,0000	–22,0000
1	2,1818	–5,6258	–9,9174	–11,0909
2	1,2876	–0,4769	–2,3987	–5,7258
3	0,8687	0,0991	–0,5265	–3,2122
4	0,7048	0,1466	–0,0806	–2,2287
5	0,6686	0,1481	–0,0039	–2,0117
6	0,6667	0,1481	0,0000	-2,0000

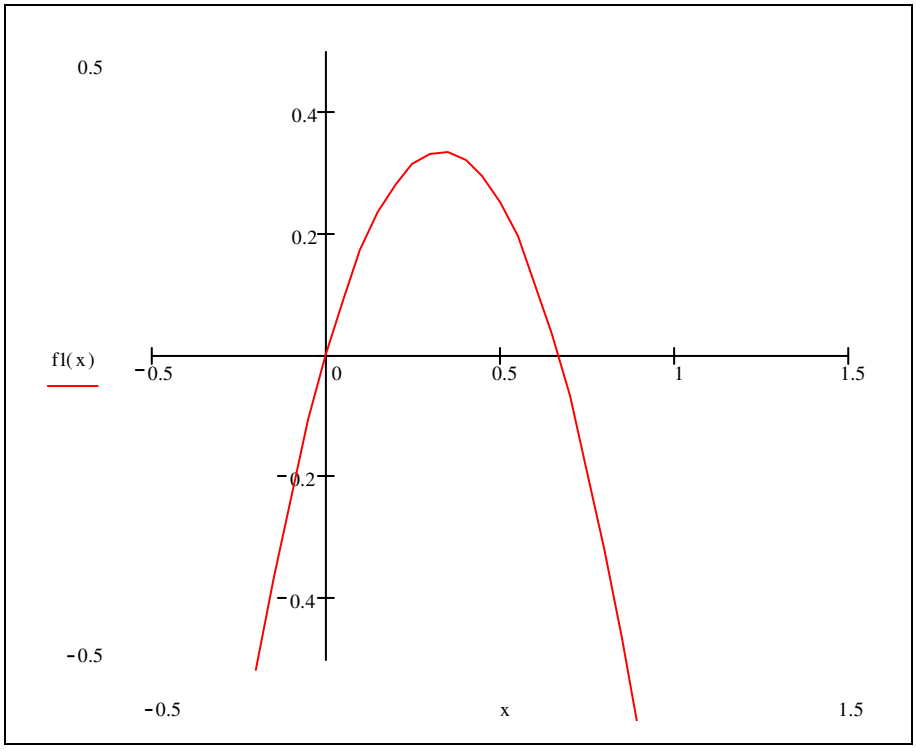
$$x^1 = x^0 - \frac{f'(x^0)}{f''(x^0)}$$
$$x^1 = 4 - \frac{-40}{-22} = 2{,}18$$

Achtung bei Newton-Raphson: Es kann leicht über das Ziel hinausgehen. Daher generell möglichst nahe bei dem vermutetem Optimum anfangen!

z.B.:

$$f(x) = -x^3 + x^2$$
$$f'(x) = -2x^2 + 2x$$


0.4
0.2
f(x)
−0.5
0
0.5
1
1.5
−0.2
−0.4


0.5
0.4
0.2
fl(x)
−0.5
0
0.5
1
1.5
−0.2
−0.4
-0.5
-0.5
x
1.5

b) Zweidimensional

Maximum gesucht von

$$f(x_1,x_2) = -\left(x_1^2 + x_2^2 + e^{x_1 - x_2}\right)$$

(Ronning 1991, S. 222 ff)

1. Ableitungen:

$$\frac{\partial f(x_1,x_2)}{\partial x_1} = -2x_1 - e^{x_1-x_2} = f_1(x)$$

$$\frac{\partial f(x_1,x_2)}{\partial x_2} = -2x_2 + e^{x_1-x_2} = f_2(x)$$

2. Ableitungen:

$$H = \frac{\partial^2 f(x_1,x_2)}{\partial x_i \partial x_j} \quad (i = 1,2)$$

Hesse'sche Matrix H = $\begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix} = \begin{pmatrix} -2 - e^{x_1-x_2} & e^{x_1-x_2} \\ e^{x_1-x_2} & -2 - e^{x_1-x_2} \end{pmatrix}$

Newton-Raphson:

$$x^1 = x^0 - H_{x^0}^{-1} g_{x^0}$$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^1 = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^0 - H^{-1} \begin{pmatrix} -2x_1 - e^{x_1-x_2} \\ -2x_2 + e^{x_1-x_2} \end{pmatrix}$$

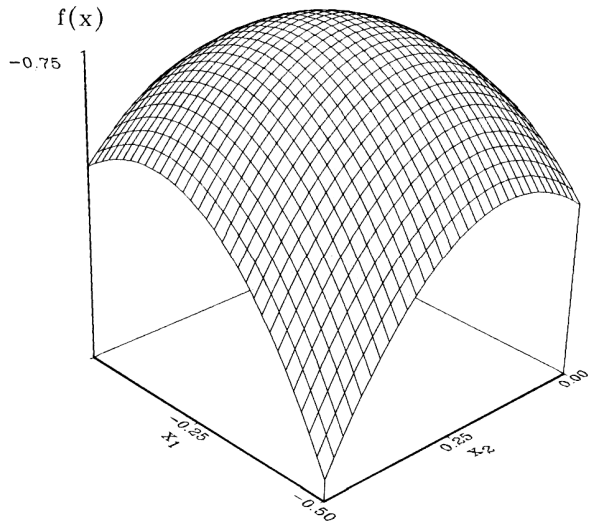
$$g_{x^0} = \begin{pmatrix} -2x_1 - e^{x_1-x_2} \\ -2x_2 + e^{x_1-x_2} \end{pmatrix}$$

g = Gradient (Vektor der ersten Ableitungen); H^{-1} = Inverse von H

Iteration	x_1	x_2	$f(x_1x_2)$	$f_1(x)$	$f_2(x)$	det H
0	2,494	–2,494	–159,0	151,6	–151,6	1618,0
1	1,98	–1,98	–60,33	56,45	–56,45	590,4
2	1,453	–1,453	–22,49	21,18	–21,18	214,0
....			....			
8	–0,2836	0,2836	–0,728	$4{,}47E^{-0.09}$	$-4{,}47E^{0.09}$	6,269
9	–0,2836	0,2836	–0,728	0,0	0,0	6,269

$$x^{0\,pt} = (-0,2836; 0,2836)$$

Graph von $f(x) = -\left(x_1^2 + x_2^2 + \exp(x_1 - x_2)\right)$ (Ronning 1991, S. 225)



Weitere Verfahren der nichtlinearen Optimierung : z.B. Greene 1990, Chapter 12

4 Hypothesentests, Modellgüte und Interpretation

4.1 Hypothesentests

Hypothesentests bezüglich der geschätzten Parameter $\hat{\beta}$

z.B. $H_o : \beta_K = 0$
$H_1 : \beta_K \neq 0$

oder Linearkombination der Parameter β_K (über $\hat{\beta}_K$).

Damit Teststatistiken gebildet werden können, sind Standardabweichungen σ_β der Koeffizienten β notwendig.

Aus der Diagonalen der asymptotischen Varianz-Kovarianzmatrix:

$$Cov\left(\hat{\beta}\right) = \left(-E\frac{\partial^2 \ln L}{\partial\beta\,\partial\beta'}\right)^{-1} = \sum\nolimits_i \left(f_i^2 \left(F_i\left(1-F_i\right)\right) x_i - x_i'\right)^{-1}$$

(Inverse des Erwartungswertes der negativen Hesse'schen Matrix) mit f_i, F_i an der Stelle $x_i\hat{\beta}$ berechnet sich der Schätzer für asymptotische Standardabweichungen $\hat{\sigma}$ aus der 2. Wurzel der Diagonalelemente von $Cov\left(\hat{\beta}\right)$.

f_i, F_i aus PROBIT, LOGIT,....

Da ML Schätzer asymptotisch normalverteilt sind, werden Hypothesentests $H_O: \beta_K = \beta^*$ über die Normalverteilung mit $z = \frac{\hat{\beta}_K - \beta^*}{\hat{\sigma}_{\hat{\beta}_K}}$ bzw. mit $t = \frac{\hat{\beta}_K - \beta^*}{\hat{\sigma}_{\hat{\beta}_K}}$ der t-Verteilung mit ν=n-k-1 Freiheitsgraden durchgeführt (siehe auch Long 1997, S. 85-101 oder Greene 1997, chapter 19,4).

4.2 Gütemaße und Spezifikationstests

Es gibt verschiedene 'measures of fit', die die Modellgüte aus der Prognosegüte der Ereignisse bewerten. Da aber das geschätzte ‚discrete choice'-Modell – im Gegensatz zum linearen Regressionsmodell, in dem b das Bestimmtheitsmaß R^2 maximiert – nicht ein Gütemaß maximiert, sind diese Maße mit Vorsicht zu genießen.

Gütemaße

- **Mc Fadden's Pseudo R^2 (oder Likelihood Ratio Index)**

$$R^2 = 1 - \left[\frac{\log L\left(\hat{\beta}^{opt}\right) \text{ im Maximum}}{\log L \text{ mit } \hat{\beta}_j = 0 \text{ außer der Konstanten } \hat{\beta}_0 \neq 0} \right] = 1 - \frac{\ln L_u}{\ln L_r}$$

mit L_u als unrestricted (volles Modell) und L_r als restricted (nur mit Konstante) Likelihood,
$R^2 \cong$ Anteil der erklärten Varianz an der Gesamtvarianz
$0 \leq R^2 \leq 1$ je näher der 1, desto besser.
Korrigiert um die Anzahl der erklärenden Variablen K (Ben-Akiva und Lerman 1985, S. 167):

$$R^2_{McFadden} = 1 - \frac{\ln\ L_u - K}{\ln\ L_r}$$

- **Korrelationskoeffizient zwischen y_i und $\hat{F}_i$**
- **Count R^2**
  aus Anzahl der korrekten Vorhersagen, mit Vorhersagen

$$\hat{y}_i = \begin{cases} 1 & \text{wenn } \hat{F}_i \geq 0{,}5 \\ 0 & \text{wenn } \hat{F}_i < 0{,}5 \end{cases} \qquad \hat{F}_i = \left(x_i \hat{\beta}\right) = P\left(y = 1 / x_i \hat{\beta}\right)$$

		Vorhersage $\hat{y}=1$	Vorhersage $\hat{y}=0$	
Beobachtung	$y=1$	n_{11}	n_{12}	$n_{1.}$
	$y=0$	n_{21}	n_{22}	$n_{2.}$
		$n_{.1}$	$n_{.2}$	$n_{..}=N$

Count R^2 (Maddala 1992, S. 304)

$$R^2_{count} = \frac{1}{N} \sum_i n_{ii} = \frac{1}{N} (n_{11} + n_{22})$$

wobei n_{ii} die korrekten Vorhersagen sind.

Korrigiert um die größte Zeilensumme

$$R^2_{count,adj} = \frac{\sum_i n_{ii} - \max(n_{.i})}{N - \max(n_{.i})}$$

- **McKelvey und Zavoina's R²**
  für LOGIT-Modelle, ist anderen Maßen überlegen: Veall und Zimmermann 1990

$$\text{Pseudo } R^2 = \frac{SQE}{SQE + n(\pi^2)/3}$$

mit

$$SQE = \sum_i \left(x_i\beta - y^*\right)^2 \quad \Big| \quad y^* = n^{-1}\sum_i x_i\beta$$

Für das MULTINOMIALE LOGIT schlägt Merz 1992 ein (gewichtetes) Mittel der einzelnen Pseudo R² vor (siehe auch Abschnitt III).

- **Maddala's Pseudo R²**

$$Pseudo\,R^2 = 1 - \left[\frac{L\left(\hat\beta_o, \beta_{k\neq 0} = 0\right)}{L\left(\hat\beta^{opt}\right)}\right]^{2/n} = 1 - \left[\frac{L_r}{L_u}\right]^{2/N}$$

$$0 \le R^2 \le 1 - L\left(\hat\beta_o\right)^{2/n}$$

Nachteil: sogar bei perfektem fit ist R² kleiner 1 (Maddala 1983, S. 39 f.).

- **Cragg und Uhler (1970) Pseudo R²**

$$Pseudo\,R^2 = \frac{L^{2/n}\left(\hat\beta^{opt}\right) - L^{2/n}\left(\hat\beta_o\right)}{1 - L^{2/n}\left(\hat\beta_o\right)} = \frac{L_u{}^{2/n} - L_r{}^{2/n}}{1 - L_r{}^{2/n}}$$

Kritische und vergleichende Diskussion: Greene 1997, chapter 19.4.2, Maddala 1983, S. 39 ff, Long 1997, S. 102-113 (mit Informations-Maßen von Akaike (AIC) nach Bayes (BIC).

Spezifikationstests

- **Gesamtansatz: Likelihood Ratio (LR) Test**

$$LR - Test = -2\left(\log L_r - \log L_u\right) = 2\left(\log L_u - \log L_r\right) = 2\log\left(\frac{L_u}{L_r}\right)$$

$L_u =$ unrestricted model $y = f\left(one, x_1, ..., x_K\right)$

$L_r =$ restricted model $y = f\left(one\right)$

$H_0:$ keine Differenz zwischen nicht-restringiertem und restringiertem Modell

H_0 wird abgelehnt, wenn LR-Test > χ^2 oder P-value < α (α = Signifikanzniveau, Chi²-Test); Anzahl der Freiheitsgrade = Anzahl der Restriktionen: $K_u - K_r$

Zwei besonders wichtige Spezifikationstests bezüglich der einzelnen unabhängigen Variablen: ausgelassene (omitted) Variablen und Heteroskedastizität (Greene 1993, S. 648; 1997, S. 889)

- **Ausgelassene (omitted) Variable (Greene 1997, S. 889)**

  $H_0 : y^* = x_1\beta_1 + \varepsilon$

  $H_1 : y^* = x_1\beta_1 + x_2\beta_2 + \varepsilon$

  Lagrange Multiplier (LM) Test: $LM = g_0' V_0^{-1} g_0$

  g_0 = Vektor der 1. Ableitungen von $\log L$ unter H_1, aber bewertet bei ML unter der Annahme H_0

  V_0^{-1} = Varianzmatrix der ML-Schätzer unter H_1 mit dem ML Schätzer unter der Annahme H_0

  auch als Wald- oder LR-Test.

- **Heteroskedastizität**
  LP: bei Heteroskedastizität immer noch unverzerrt und konsistent, jedoch ineffizient
  LOGIT/PROBIT: ′pessimistischer′
  LM-, LR- oder Wald-Test: vgl. Greene 1993, S. 649 f., 1997, S. 889f.

4.3 Interpretation, Ableitungen und Elastizitäten

Der Einfluss der unabhängigen Variablen auf die abhängige Variable wird im linearen Fall durch den Regressionskoeffizienten gemessen. Bei den nichtlinearen Modellen Logit und Probit sind zusätzliche Berechnungen notwendig (siehe auch das Kreditvergabebeispiel in South Carolina von vorne) um aus dem nichtlinearen Ansatz auf die Wahrscheinlichkeiten zu schließen..

Die ersten Ableitungen geben ∂y bei „Änderung“ von ∂x an. Unter Berücksichtigung der nichtlinearen Wahrscheinlichkeitstransformation gilt:

Marginale Effekte: Einfluss von x_{ik} auf die Wahrscheinlichkeit $P(y_i = 1)$

LP $$\frac{\partial}{\partial x_{ik}}(x_i\beta) = \beta_k$$

LOGIT $$\frac{\partial}{\partial x_{ik}}\Lambda(x_i\beta) = \frac{\exp(x_i\beta)}{\left[1+\exp(x_i\beta)\right]^2}\beta_k$$

PROBIT $$\frac{\partial}{\partial x_{ik}}\Phi(x_i\beta) = \phi(x_i\beta)\beta_k$$

bzw. $$\frac{\partial P_i}{\partial x_i} = \beta f_i = \begin{cases} \beta & \text{LP} \\ \beta\, p_i(1-p_i) & \text{LOGIT} \\ \beta\, f\left[F^{-1}(p_i)\right] & \text{PROBIT} \end{cases}$$

Größter Effekt bei $P_i = 0{,}5,$ nähert sich Null, wenn P_i sich Null oder Eins nähert.

Elastizitäten: Relative Änderung von y bei relativer Änderung von x_k

$$\varepsilon_{zx} = \frac{\partial z}{z} \Big/ \frac{\partial x_k}{x_k} = \frac{\partial z}{\partial x_k} \Big/ \frac{z}{x} = \frac{\partial \ln z}{\partial \ln x_k}$$

$|\varepsilon| > 1$ elastisch
$|\varepsilon| < 1$ unelastisch

Für die diskreten Entscheidungsmodelle ist für z dann die jeweilige Wahrscheinlichkeit P (z=1) einzusetzen.

Aussage: Eine 1%-ige Änderung von x_i bewirkt eine ε-ige Änderung der erklärten Variablen z, wobei z (=F(x_iß)) die jeweilige Wahrscheinlichkeit des LP-, des Logit- oder des Probit-Modells ist.

4.4 Stata, LIMDEP und GAUSS: Gütemaße, Spezifikationstests

Stata:	Likelihood Ratio Test
LIMDEP:	PROBIT ML-Schätzung, Goodness-of-fit Maße
GAUSS:	Pseudo R^2 für PROBIT-, LOGIT-, Extremwertmodell
LIMDEP:	Pseudo R^2 nach McKelvey und Zavoina
LIMDEP:	Modellgüte, LR Test
LIMDEP:	Spezifikationstest, Homoskedastizität

Stata: Likelihood Ratio Test

Program evaluation example data, "National Supported Work (NSW) Demonstration"
Training program for men with poor labor market histories (from the1970s)

jtrain3.dta
http://qcpages.qc.cuny.edu/~rvesselinov/statadata/jtrain3.dta

```
  train (1 or 0): treatment with a job training program (began two years prior to
    1978)
  re78 response: real labor market earnings in 1978
  age
  educ: years of education
  black (1 or 0), hispanic (1 or0)
  maritial status
  re74 and re75: real earnings two years prior to the start of the program etc.
```

```
------------------------------------------------------------------------------
. use jtrain3.dta, clear
. * model_1
. probit train age educ
Iteration 0:   log likelihood = -672.64954
Iteration 1:   log likelihood = -551.12652
Iteration 2:   log likelihood = -536.20456
Iteration 3:   log likelihood = -536.09612
Iteration 4:   log likelihood = -536.09609
Probit regression                                 Number of obs   =       2675
                                                  LR chi2(2)      =     273.11
                                                  Prob > chi2     =     0.0000
Log likelihood = -536.09609                       Pseudo R2       =     0.2030
------------------------------------------------------------------------------
       train |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |  -.0686465   .0056077   -12.24   0.000    -.0796374   -.0576557
        educ |  -.1774961   .0173217   -10.25   0.000    -.2114459   -.1435463
       _cons |   2.591852   .2928778     8.85   0.000     2.017822    3.165882
------------------------------------------------------------------------------
. eststo model_1
```

LR-Test
L (restricted): train=f(_cons)
L (unrestricted): train=f(age, educ, _cons)

Kritischer chi2-Wert (df=2, Anzahl der zusätzlichen Variablen)

Da LR chi2 = 273.11 > chi2 crit = 13.816; oder p-value < 0,1%

=> H0: kein Unterschied, ablehnen, d.h. model_1 ist signifikant besser als das restringierte Modell.

```
. * model_2
. probit train age educ black hisp
Iteration 0:   log likelihood = -672.64954
Iteration 1:   log likelihood = -482.61563
Iteration 2:   log likelihood = -448.91264
Iteration 3:   log likelihood = -447.99738
Iteration 4:   log likelihood = -447.99493
Iteration 5:   log likelihood = -447.99493
```

```
Probit regression                                 Number of obs   =       2675
                                                  LR chi2(4)      =     449.31
                                                  Prob > chi2     =     0.0000
Log likelihood = -447.99493                       Pseudo R2       =     0.3340
------------------------------------------------------------------------------
       train |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |  -.0661692   .0062066   -10.66   0.000    -.0783339   -.0540044
        educ |  -.1140225   .0201033    -5.67   0.000    -.1534243   -.0746208
       black |   1.345078   .1150059    11.70   0.000     1.11967    1.570485
        hisp |   .9881963   .2176798     4.54   0.000     .5615517    1.414841
       _cons |   .9944103   .3476386     2.86   0.004     .3130512    1.675769
------------------------------------------------------------------------------
. eststo model_2
```

Modellvergleich

LR-test=2(logL(model_2) – logL(model_1))
=2(449.31 – 273.11) = 352.40.

Da LR chi2 = 352.40 > chi2 crit = 13.816; oder p-value < 0,1%

=> H0: kein Unterschied, ablehnen, d.h. model_2 ist signifikant besser als das restringierte Modell model_1.

```
. * linear
. regress train age educ black hisp

      Source |       SS       df       MS              Number of obs =    2675
-------------+------------------------------           F(  4,  2670) =  122.21
       Model |  26.6494742     4  6.66236856            Prob > F      =  0.0000
    Residual |  145.556133  2670  .054515406           R-squared     =  0.1548
-------------+------------------------------           Adj R-squared =  0.1535
       Total |  172.205607  2674  .064400003           Root MSE      =  .23349

------------------------------------------------------------------------------
       train |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |  -.0049167   .0004458   -11.03   0.000    -.0057909   -.0040425
        educ |  -.0067313   .0016372    -4.11   0.000    -.0099416    -.003521
       black |   .1619288    .010896    14.86   0.000     .1405633    .1832942
        hisp |   .0909079    .025232     3.60   0.000     .0414317     .140384
       _cons |   .2678323   .0294746     9.09   0.000      .210037    .3256276
------------------------------------------------------------------------------

. eststo linear
. esttab model_1 model_2 linear,nogaps p star wide nopar

--------------------------------------------------------------------------------------------
                    (1)                         (2)                         (3)
                  train                       train                       train
--------------------------------------------------------------------------------------------
main
age             -0.0686***       0.000      -0.0662***       0.000     -0.00492***       0.000
educ             -0.177***       0.000       -0.114***       0.000     -0.00673***       0.000
black                                         1.345***       0.000        0.162***       0.000
hisp                                          0.988***       0.000       0.0909***       0.000
_cons             2.592***       0.000        0.994**        0.004        0.268***       0.000
--------------------------------------------------------------------------------------------
N                    2675                        2675                        2675
--------------------------------------------------------------------------------------------
p-values in second column
* p<0.05, ** p<0.01, *** p<0.001
```

LIMDEP: PROBIT ML-Schätzung, Goodness-of-fit Maße

Wahrscheinlichkeit von Studenten für Studiengebühren zu sein - Studierendenumfrage aus der Statistik-Vorlesung (Merz) mit Informationen zu Studiengebühren 2009

```
Normal exit from iterations. Exit status=0.
+--------------------------------------------+
| Binomial Probit Model                      |
| Maximum Likelihood Estimates               |
| Model estimated: May 20, 2008 at 09:25:11AM.|
| Dependent variable               PROSTG    |
| Weighting variable                 None    |
| Number of observations              140    |
| Iterations completed                  6    |
| Log likelihood function       -57.51840    |
| Restricted log likelihood     -60.88622    |
| Chi squared                    6.735646    |
| Degrees of freedom                    6    |
| Prob[ChiSqd > value] =         .3459879    |
| Hosmer-Lemeshow chi-squared =   4.32216    |
| P-value=  .82695 with deg.fr. =       8    |
+--------------------------------------------+
+---------+-------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+-------------+----------------+--------+---------+----------+
          Index function for probability
 Constant  .1150980712E-02      1.6322285     .001   .9994
 MANN         .3653463677       .29101900    1.255   .2093     .44285714
 ALTER    -.8990327285E-01  .77731876E-01   -1.157   .2474     21.814286
 SEM       .5928557516E-01       .30598104     .194   .8464     2.1428571
 JOB          .2784881887       .28041722     .993   .3207     .50714286
 BWL          .5884789523       .70455606     .835   .4036     .57142857
 WISO         .6279986423       .46319604    1.356   .1752     .18571429
 (Note: E+nn or E-nn means multiply by 10 to + or -nn power.)
+----------------------------------------+
| Fit Measures for Binomial Choice Model |
| Probit   model for variable PROSTG     |
+----------------------------------------+
| Proportions P0= .842857   P1= .157143  |
| N =     140 N0=     118   N1=      22  |
| LogL =   -57.51840 LogL0 =   -60.8862  |
| Estrella = 1-(L/L0)^(-2L0/n) = .04829  |
+----------------------------------------+
|     Efron |  McFadden  |  Ben./Lerman  |
|    .04433 |    .05531  |       .74687  |
|    Cramer | Veall/Zim. |     Rsqrd_ML  |
|    .04457 |    .09868  |       .04697  |
+----------------------------------------+
| Information  Akaike I.C. Schwarz I.C.  |
| Criteria         .92169     149.62829  |
+----------------------------------------+
Frequencies of actual & predicted outcomes
Predicted outcome has maximum probability.
Threshold value for predicting Y=1 = .5000
           Predicted
------  ----------  +  -----
Actual       0    1  |  Total
------  ----------  +  -----
  0        118    0  |    118
  1         22    0  |     22
------  ----------  +  -----
Total      140    0  |    140
```

GAUSS: Pseudo R² für PROBIT, LOGIT, Extremwertmodell

Prozeduren zur Berechnung von Pseudo-R^2-Maßzahlen für binäre Entscheidungsmodelle (PROBIT, LOGIT, Extremwertmodelle) (Hilmar Schneider)

```
Die im folgenden als Call-Argumente verwendeten Kürzel bedeuten:

mll:    mittlere Log-Likelihood
n:      Gesamtzahl aller Beobachtungen
c:      Summe der Beobachtungen, für die gilt y = 1

file:   Filename-String
ix:     Spaltenvektor mit den Spaltenindices der erklärenden Variablen
b:      Spaltenvektor der Schätzparameter korrespondierend zu ix
modell: Modelltyp-String (zulässige Angaben: "logit" oder "probit")

PROC mcfadden(mll,n,c);              @ McFadden's Pseudo R²                @
   LOCAL mll, n, c, null;
   null = c*LN(c/(n-c)) - n*LN(n/(n-c));
   RETP( 1-mll*n/null );
ENDP;

PROC maddala(mll,n,c);               @ Maddala's Pseudo R²                 @
   LOCAL mll, n, c, null;
   null = c*LN(c/(n-c)) - n*LN(n/(n-c));
   RETP( 1-EXP(2*(null/n-mll)) );
ENDP;

PROC cragguhl(mll,n,c);              @ Cragg/Uhler's Pseudo R²             @
   LOCAL mll, n, c, null;
   null = c*LN(c/(n-c)) - n*LN(n/(n-c));
   RETP( (1-EXP(2*(null/n-mll)))/(1-EXP(2*null/n)) );
ENDP;

PROC mz(file,ix,b,modell);           @ McKelvey/Zavoina's Pseudo R²        @
   LOCAL file, ix, b, modell, fh, dta, maxrecs, n, x, sy, sqe;

   sy      = 0;
   sqe     = 0;

   OPEN fh = ^file;
   n       = ROWSF(fh);
   maxrecs = TRUNC(8190/COLSF(fh));
   DO UNTIL EOF(fh);
      dta = READR(fh,maxrecs);
      x   = ONES(ROWS(dta),1)~dta[.,ix];
      sy  = sy + SUMC(x*b);
      sqe = sqe + b'x'x*b;
   ENDO;
   fh  = CLOSE(fh);

   sqe = sqe - sy*sy/n;

   IF     UPPER(modell) $== "LOGIT";
      RETP( sqe/(sqe + n*PI^2/3) );
   ELSEIF UPPER(modell) $== "PROBIT";
      RETP( sqe/(sqe + n) );
   ENDIF;

ENDP;
```

$$Sy: \quad \sum_i \hat{y}_i = \sum_i \underline{x}_i \hat{\underline{\beta}}$$

$$Sqe: \quad \sum_i (\hat{y}_i - \bar{y})^2 = \sum_i \hat{y}_i^2 - n\bar{y}^2$$

$$= \sum_i \hat{y}_i^2 - \frac{\sum_i \hat{y}_i \cdot \sum_i \hat{y}_i}{n}$$

$$= \sum_i \left(\underline{x}_i \hat{\underline{\beta}}\right)^2 - \frac{\left(\sum_i \hat{y}_i\right)^2}{n}$$

$$= \hat{\underline{\beta}}' \underline{x}' \underline{x} \hat{\underline{\beta}} - \frac{Sy \cdot Sy}{n}$$

Pseudo R² nach McKelvey und Zavoina LIMDEP: PROBIT. LIM

```
?=====================================================================
?  This program computes a Pseudo R² for a probit model based on the
?  formula given by McKelvey and Zavoina in their paper on the ordered
?  probit model.  The formula suggested is
?
?         E[y*|y]  =  yf  =  ß'x  +  lambda
?
?         R² = N * [ var(yf) / (1 + var(yf) ]
?
?  where  lambda is the inverse Mill's ratio usually kept for SELECT.
?  (To get the MZ result exactly, we'd use 1/N not 1/(N-1) in var(yf).)
? User must define the data matrix with the following line:          =
?                                                                    =
?        NAMELIST          ; X = ... $                             =
?                                                                    =
? User's dependent variable is Y.  Change the following line        =
?                                                                    =
?        CREATE            ; Y = the dependent variable $          =
?
?=====================================================================
   PROBIT    ; lhs = y ; rhs = X ; hold(imr=lambda) $
   CREATE    ; yf = dot(x,b) + lambda $
   CALC      ; zm = (n-1)*var(yf) / (n + (n-1)*var(yf)) $
```

Modellgüte, Likelihood-Ratio Test LIMDEP: PROBITLR. LIM

```
?=====================================================================
? Likelihood ratio test                                              =
? univariate probit model                                            =
?                                                                    =
? reject the null hypothesis                                         =
?  (H0: no difference between unrestricted and restricted  model)    =
?  if                                                                =
?   lrtest > tabchi95 or tabchi99, or equivalent                     =
?   pvalue < 5% or 1% significance level          (corrected JM      =
?=====================================================================


type; restricted case$
type; ---------------$

probit; lhs=y; rhs=one;
calc; lr = logl$


type; unrestricted case$
type; -----------------$

PROBIT     ; Lhs = y ; Rhs = X $
calc; lu = logl$

type; Likelihood ratio test:$
type; ----------------------$
calc; lrtest = 2*(lu-lr)
    ; tabchi95=ctb(.95,1)
    ; tabchi99=ctb(.99,1)
    ; pvalue = 1-chi(lrtest,1)$

type; Likelihood ratio test: reject H0$
type; (H0: no difference between unrestricted (Probit y=one)$
type; and restricted model (Probit y=x) if$
type;   lrtest > tabchi95 or tabchi99,  or equivalent$
type;   pvalue < 5% or 1% significance level$
```

Spezifikationstest, Homoskedastizität LIMDEP: PROBIT.LM LIM

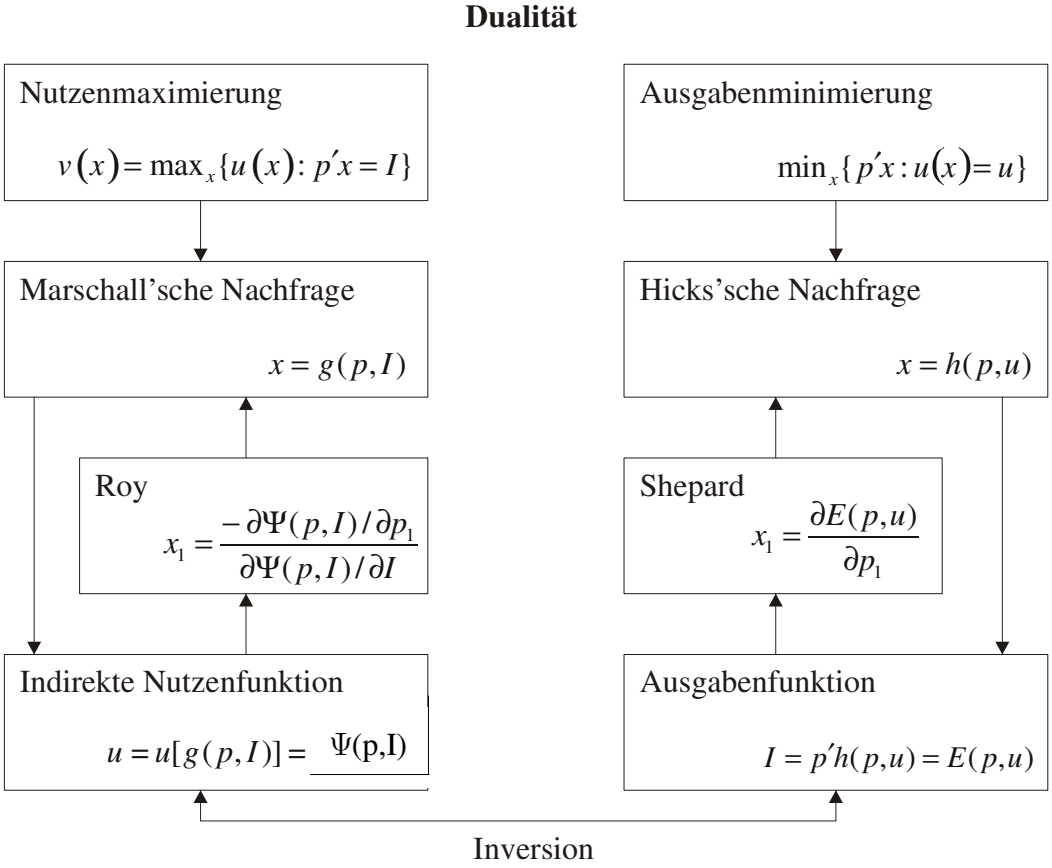
```
?=================================================================
? Lagrange multiplier test for homoskedasticity in the           =
? univariate probit model                                        =
?                                                                =
? User must define the data matrix with the following line:      =
?                                                                =
?        NAMELIST          ; X = ... $                           =
?                                                                =
? User must define the variables in the heteroskedasticity model =
? Z must not contain ONE                                         =
?                                                                =
?        NAMELIST          ; Z = ... $                           =
?                                                                =
? User's dependent variable is Y.  Change the following line     =
?                                                                =
?        CREATE            ; Y = the dependent variable $        =
?=================================================================
PROBIT      ; Lhs = y ; Rhs = X $
CREATE      ; qi=2 * y - 1
            ; xb = dot(X,b)
            ; vi = -qi * lmm( qi * xb )
            ; vb = vi ; vg = -vi * xb$
MATRIX      ; gb = Xdot(X,one,vb)
            ; gg = Xdot(Z,one,vg)
            ; g  = gb / gg
            ; h  = BHHH(X,Z,vb,vg)
            ; LMSTAT = g' | sinv(H) | g $
```

5 Nutzentheoretische Grundlagen und Binäres Entscheidungsmodell

5.1 Mikroökonomische Entscheidungstheorie

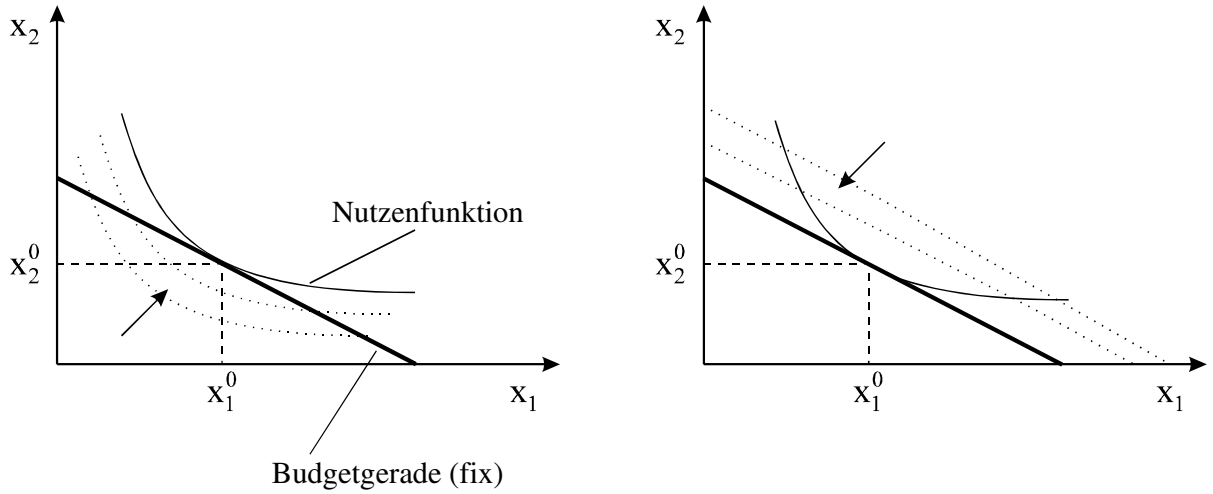
Mikroökonomisch nachfrageorientiertes Modell

z.B. Merz 2012, S. 18; Deaton und Muellbauer 1980, 'Standard'-textbook



Fazit: Optimale Entscheidung nach nachzufragender Menge

$x_i = f(p, I(s))$ ist eine kontinuierliche Größe (s = sozioökonomische Charakteristika).



Primales Problem **Duales Problem**

Optimum: Primales Problem: Tangentialpunkt Budgetgerade
Duales Problem: Tangentialpunkt Kostengerade

Gleiche Vorgehensweise auch bei ‚Kinked budget curves' (Steigung der Budgetgerade (=Preisverhältnis) ändert sich in Abhängigkeit von den Mengen, z.B. bei Rabatten).

5.2 Mikroökonomisches Beispiel zur Dualität

Beispiele zur Illustration des mikroökonomischen Entscheidungsmodells:

$$u = v(x) = x_1 x_2$$
$$p'x = p_1 x_1 + p_2 x_2 = I$$

p_i = Preise
x_i = Mengen
I = Einkommen
u = Nutzen

Primales Problem: Nutzenmaximierung $v(x) = \max_x \{u(x) : p' \cdot x = I\}$

$$L(x,\lambda) = u(x) - \lambda(p'x - I)$$
$$L(x,\lambda) = x_1 x_2 - \lambda(p_1 x_1 + p_2 x_2 - I)$$

$$(1)\ \frac{\partial L}{\partial x_1} = x_2 - \lambda \cdot p_1 = 0 \quad \rightarrow\ \lambda = \frac{x_2}{p_1} = \frac{\partial u / \partial x_1}{p_i} = \frac{\partial u}{\partial (p_i \cdot x_i)} \rightarrow$$

$$(2)\ \frac{\partial L}{\partial x_2} = x_1 - \lambda \cdot p_2 = 0 \quad \rightarrow\ \lambda = \frac{x_1}{p_2} \qquad \lambda = \textit{Grenznutzen des Einkommens}$$

$$(3)\ \frac{\partial L}{\partial \lambda} = p_1 \cdot x_1 + p_2 \cdot x_2 - I = 0$$

$$\Rightarrow x_2 / x_1 = p_1 / p_2; \quad x_2 = \frac{p_1}{p_2} x_1; \quad x_1 = \frac{p_2}{p_1} x_2$$

$$\text{in } (3) \qquad p_1 x_1 + p_2 \frac{p_1}{p_2} x_1 = I$$

$$2 p_1 x_1 = I$$

$$x_1 = \frac{I}{2 p_1}$$

$\Rightarrow$ Marshall'sche Nachfrage

$$x_i = \frac{I}{2 p_i} = g_i(p, I)$$

Indirekte Nutzenfunktion

$$u = v\left[g(p, I) \right] = v\left[\frac{I}{2 \cdot p_1}, \frac{I}{2 \cdot p_2} \right] = \Psi(p, I)$$

$$u = x_1 \cdot x_2 = \frac{I}{2 \cdot p_1} \cdot \frac{I}{2 \cdot p_2} = \frac{I^2}{4 \cdot p_1 \cdot p_2}$$

$$\Psi(p, I) = \frac{I^2}{4 \cdot p_1 \cdot p_2}$$

Roy's Indentität → Marshall'sche Nachfrage

$$x_i = \frac{-\partial \Psi(p, I) / \partial p_i}{\partial \Psi(p, I) / \partial I}$$

$$= \frac{-\left(I^2 / 4 \cdot p_1 \cdot p_2 \right) / \partial p_i}{\left(I^2 / 4 \cdot p_1 \cdot p_2 \right) / \partial I}$$

$$y = I^2 / \overbrace{\left(4 \cdot p_1 \cdot p_2 \right)}^{z} = I^2 \cdot z^{-1}$$

$$\frac{\partial y}{\partial p_1} = \frac{\partial y}{\partial z} \cdot \frac{\partial z}{\partial p_1} = -I^2 \cdot z^{-2} \cdot 4 \cdot p_2$$

$$= \frac{-I^2}{\left(4 \cdot p_1 \cdot p_2 \right)^2} \cdot 4 \cdot p_2 = \frac{-I^2}{4 \cdot p_1^2 \cdot p_2}$$

$$\frac{\partial y}{\partial p_2} = -I^2 \cdot z^{-2} \cdot 4 \cdot p_1$$

$$= \frac{-I^2}{4 \cdot p_1 \cdot p_2^2}$$

$$\frac{\partial \left(I^2 / 4 \cdot p_1 \cdot p_2 \right)}{\partial I} = \frac{2 \cdot I}{4 \cdot p_1 \cdot p_2}$$

$$x_1 = \frac{-\left[-I^2/4 \cdot p_1^2 \cdot p_2\right]}{2 \cdot I/4 \cdot p_1 \cdot p_2} = \frac{I^2 \cdot 4 \cdot p_1 \cdot p_2}{2 \cdot I \cdot 4 \cdot p_1^2 \cdot p_2}$$
$$= I/(2 \cdot p_1)$$

$$x_2 = \frac{-\left[-I^2/4 \cdot p_1 \cdot p_2^2\right]}{2 \cdot I/4 \cdot p_1 \cdot p_2} = \frac{I^2 \cdot 4 \cdot p_1 \cdot p_2}{2 \cdot I \cdot 4 \cdot p_1 \cdot p_2^2}$$
$$= I/(2 \cdot p_2)$$
$$x_i = I/(2 \cdot p_i) = g(p, I)$$

Duales Problem: Ausgabenminimierung $E(p,u) \equiv \min\{p'x : v(x) = u\}$

$$L(x,\lambda) = p' \cdot x - \lambda\left(v(x) - u\right)$$
$$v(x) = x_1 \cdot x_2$$

$$(1^*) \quad \frac{\partial L}{\partial x_1} = p_1 - \lambda x_2 = 0 \Rightarrow \lambda = \frac{p_1}{x_2}$$

$$(2^*) \quad \frac{\partial L}{\partial x_2} = p_2 - \lambda x_1 = 0 \Rightarrow \lambda = \frac{p_2}{x_1}$$

$$(3^*) \quad \frac{\partial L}{\partial \lambda} = x_1 x_2 - u = 0$$

$$\Rightarrow \frac{p_1}{x_2} = \frac{p_2}{x_1} \Rightarrow x_1 = \frac{p_2}{p_1} x_2; x_2 = \frac{p_1}{p_2} x_1$$

in (3^*)

$$\frac{p_2}{p_1} = x_2 \cdot x_1 - u = 0$$
$$x_2 = \left(\frac{p_1}{p_2} \cdot u\right)^{1/2}$$

bzw.

$$x_1 \frac{p_1}{p_2} x_1 - u = 0$$
$$x_1 = \left(\frac{p_2}{p_1} u\right)^{1/2}$$

Hicks'sche Nachfrage

$$x_i = \left(\frac{p_j}{p_i} u\right)^{1/2} = h(p,u) \qquad i, j = 1, 2$$

Ausgabenfunktion ('cost function')

$$I = p'h(p,u) = E(p,u)$$

$$I = p_1\left(\frac{p_2}{p_1}u\right)^{1/2} + p_2\left(\frac{p_1}{p_2}u\right)^{1/2}$$

$$I = \sqrt{p_2 p_1 u} + \sqrt{p_1 p_2 u}$$

$$I = 2(p_1 p_2 u)^{1/2}$$

Sheppard's Lemma → Hicks'sche Nachfrage

$$x_i = h(p,u) = \frac{\partial E(p,u)}{\partial p_i}$$

$$x_1 = \frac{2\cdot(p_1\cdot p_2\cdot u)^{1/2}}{\partial p_1} = \frac{2\cdot p_1^{1/2}\cdot p_2^{1/2}\cdot u^{1/2}}{\partial p_1} = p_1^{-1/2}\cdot p_2^{1/2}\cdot u^{1/2} = \left(\frac{p_2}{p_1}\cdot u\right)^{1/2}$$

$$x_i = \left(\frac{p_j}{p_i}\cdot u\right)^{1/2}$$

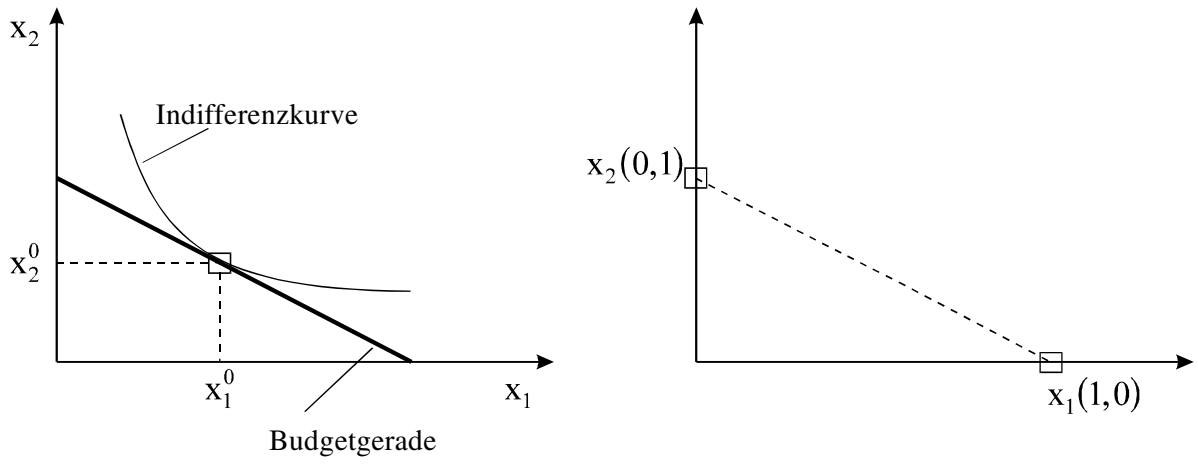
Indirekte Nutzenfunktion (Inversion der Ausgabenfuntion)

$$I^2 = 4p_1 p_2 u$$

$$u = \frac{I^2}{4p_1 p_2} = \Psi(p,I)$$

5.3 Mikroökonomische Nutzenmaximierung und stochastische Nutzenfunktion (RUM: Random Utility Maximization)

Der mikroökonomische marginalanalytische Ansatz stößt bei Problemen mit diskreten voneinander unabhängigen Alternativen schnell an seine Grenzen, da eine kleine Änderung einer Entscheidungsvariablen keine diskret voneinander zu unterscheidende Ereignisse bewirkt.



$p'x = I \qquad \Rightarrow y = p_1 x_1 + p_2 x_2$

$x_2 = \frac{y}{p_2} - \frac{p_1}{p_2} x_1$

entweder	x_1 *ja*	x_2 nein
oder	x_1 *nein*	x_2 *ja*

Kontinuierliche Entscheidung:

$x^0 = \left(x_1^0, x_2^0\right)'$

Diskrete Entscheidung:

$x_i = \begin{cases} 1 & \text{ja} \\ 0 & \text{nein} \end{cases}$

Mikroökonomisches Modell und indirekte Nutzenfunktion

Direkte Nutzenfunktion = f(Quantität | Qualität der Güter)
Indirekte Nutzenfunktion = f(Preise, Einkommen) (also Eigenschaften der Güter etc.)

falls eine konkrete indirekte Nutzenfunktion spezifiziert ist, wäre (→ Roy, Shephard) die Marshall'sche bzw. Hicks'sche Nachfragefunktion relativ einfach (Ableitung) bestimmbar.

RUM: Random Utility Maximization

Wie bereits angesprochen, wählt ein rationaler Entscheider im stochastischen Nutzenmodell RUM die Alternative, die ihm den höchsten Nutzen stiftet. Der Nutzen selbst ist nicht beobachtbar (latent), nur das eingetretene Ereignis ist beobachtbar. Der entsprechende Indikator ist dann 1 wenn $u_1 > u_2$ und 0, wenn $u_1 \leq u_2$, wenn also die Wahl auf die Miete fällt.

Lineare stochastische Nutzenspezifikation

$u_{i1} = v_{i1} + \varepsilon_{i1} = x_i \beta_1 + \varepsilon_{i1} \qquad u_{i2} = v_{i2} + \varepsilon_{i2} = x_i \beta_2 + \varepsilon_{i2}$

v_{ij} deterministische Komponente

ε_{ij} additiver Zufallsterm

Entscheider wählt die Alternative j aus allen m=2 Alternativen, die ihm den höchsten Nutzen stiftet:

$u_{ij} = \max(u_{i1}, u_{i2})$ bzw.

$y_i = \begin{cases} 1 & \textit{wenn} \quad u_1 > u_2 \ \textit{bzw.} \ u_1 - u_2 > 0 \\ 0 & \textit{sonst} \end{cases}$

$\text{Prob}\left(y_i = 1\right) = \text{Prob}\left(Z \in \text{best. Nutzendifferenzen}\right)$

oder nach Umformung

$v_{i1} + \varepsilon_{i1} > v_{i2} + \varepsilon_{i2}$ bzw. $\left(v_{i1} - v_{i2}\right) > \left(\varepsilon_{i2} - \varepsilon_{i1}\right)$ $\left(\textit{für} \text{ Alternative } 1\right)$.

Die Nutzendifferenz ist eine kontinuierliche Zufallsvariable, die auch als latente Größe im Index Funktionsmodell interpretiert werden kann, die, wenn sie über einen Schwellenwert hinausgeht, über die Indexfunktion den Eintritt eines beobachtbaren Ereignisses *j* bestimmt.

RUM-Modell und indirekte Nutzenfunktion

Die RUM-Nutzenfunktion $u_i = x_i\beta + \varepsilon_i$ kann als eine indirekte Nutzenfunktion interpretiert werden, da der Nutzen als eine Funktion der Eigenschaften der Alternative/Individuum ist.

Beispiel: Entscheidung zwischen zwei Alternativen ($m = 2$) :

Sei G die (kumulative) Verteilungsfunktion (cdf) der Differenzen $\varepsilon_{iK} - \varepsilon_{ij} = \mu_{kj} = \varepsilon_{i2} - \varepsilon_{i1}$

Wahrscheinlichkeit für $j = 1$

$$P_1 = \text{Prob}[\varepsilon_{i2} - \varepsilon_{i1} < \upsilon_{i1} - \upsilon_{i2}] = G[\upsilon_{i1} - \upsilon_{i2}]$$

Annahmen:

- υ ist lineare Funktion

  $\upsilon_{ij} = \alpha' x_j + \beta_j' y_i + \gamma' z_{ij} = \beta' x_{ji}$ ('neues' x und β)

  $-\varepsilon_{ij} = \varepsilon_j$, keine Funktion von x_{ij}, ε_1 und ε_2 statistisch unabhängig

Wahrscheinlichkeitsverteilung bezüglich υ_{ij} :

PROBIT:

G = cdf der Standardnormal-Verteilung

$$\begin{aligned} P_1 &= \Phi\left(\beta' x_{i1} - \beta' x_{i2}\right) \\ &= \Phi\left(\beta'\left(x_{i1} - x_{i2}\right)\right) \end{aligned}$$

LOGIT:

G = cdf logistisch verteilt

$$P_1 = \frac{1}{1 + \exp\left[-\left(\beta' x_{i1} - \beta' x_{i2}\right)\right]}$$

Wahrscheinlichkeitsverteilung bezüglich ε_1 und ε_2 :

PROBIT:

ist $\varepsilon_1 und \varepsilon_2$ unabhängig normal verteilt $\Rightarrow$ Differenz $\boldsymbol{\varepsilon_2 - \varepsilon_1}$ ist normalverteilt

$\rightarrow$ G ist cdf normal $\Rightarrow$ PROBIT

LOGIT:

ist $\varepsilon_1 und \varepsilon_2$ unabhängig Weibull-verteilt

$\rightarrow$ G (Differenz $\varepsilon_2 - \varepsilon_1$) $\Rightarrow$ LOGIT

Weibull-Verteilung:

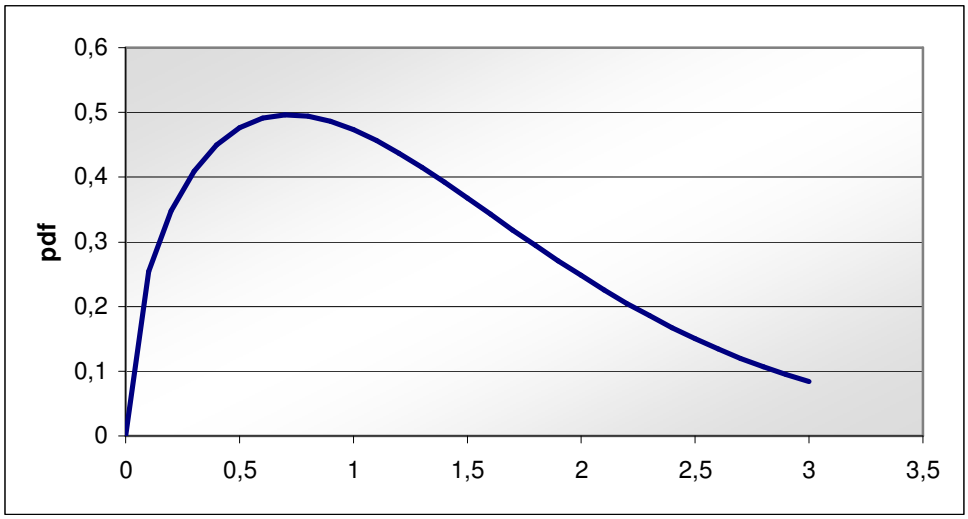
$$pdf(\varepsilon) = e^{-(\varepsilon+a)} \cdot \exp\left(-e^{-(\varepsilon+a)}\right)$$
$$cdf: \text{Prob}(\varepsilon_j < \varepsilon) = \exp\left(-e^{-(\varepsilon+a)}\right)$$

Charakteristika der Weibull-Verteilung:

$a = 0$ $\rightarrow$ Modus der Dichtefunktion = 0, Mittelwert = 0,575, Varianz = 1,622

Im Vergleich zur normalen Dichte (Mittel = 0,5, Var = 1) hat die Weibull-Verteilung ein dickeres rechtes 'Ende' und kürzeres linkes 'Ende'.

Auf die Weibull-Verteilung kommen wir bei multiplen Entscheidungen noch zurück.



6 Beispiele

6.1 LOGIT Schätzung mit stata – Video example

Stata bietet via internet Training Videos auch für binäre Regressionsmodelle an:

www.stata.com

-> Training -> Video tutorials -> Binary outcomes ->

Logistic regression in Stata, part 1: Binary predictions

(www.youtube.com/watch?v=rSU1L3-xRk0)

6.2 Training – LOGIT und PROBIT Schätzung mit stata

Program evaluation example data, "National Supported Work NSW) Demonstration"
Training program for men with poor labor market histories (from the1970s)

jtrain3.dta
http://qcpages.qc.cuny.edu/~rvesselinov/statadata/jtrain3.dta

```
  train (1 or 0): treatment with a job training program (began two years prior to
    1978)
  re78 response: real labor market earnings in 1978
  age
  educ: years of education
  black (1 or 0), hispanic (1 or0)
  maritial status
  re74 and re75: real earnings two years prior to the start of the program etc.


-------------------------------------------------------------------------------------------------------------------
name:  <unnamed>
log:  N:\Gruppen\Lehre\_Ewf\_Discrete Choice Mikrooekonometrie\Stata Examples\jtrain
version 10\logit-probit.log
  log type:  text
 opened on:  10 Apr 2014, 15:07:37
. use jtrain3.dta, clear

. * logit-1
. logit train age educ black hisp

Iteration 0:   log likelihood = -672.64954
Iteration 1:   log likelihood = -548.86008
Iteration 2:   log likelihood = -451.22767
Iteration 3:   log likelihood = -445.54543
Iteration 4:   log likelihood = -445.48077
Iteration 5:   log likelihood = -445.48076

Logistic regression                               Number of obs   =       2675
                                                  LR chi2(4)      =     454.34
                                                  Prob > chi2     =     0.0000
Log likelihood = -445.48076                       Pseudo R2       =     0.3377

------------------------------------------------------------------------------
       train |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |  -.1320748   .0126665   -10.43   0.000    -.1569008   -.1072488
        educ |  -.2292395   .0386296    -5.93   0.000    -.3049521   -.1535268
       black |    2.74895   .2621999    10.48   0.000     2.235048    3.262852
        hisp |   2.177318   .4222033     5.16   0.000     1.349815    3.004822
       _cons |   2.136474   .7024242     3.04   0.002     .7597477      3.5132
------------------------------------------------------------------------------
```

```
. eststo logit_1

.
. * probit-1
. probit train age educ black hisp

Iteration 0:   log likelihood = -672.64954
Iteration 1:   log likelihood = -482.61563
Iteration 2:   log likelihood = -448.91264
Iteration 3:   log likelihood = -447.99738
Iteration 4:   log likelihood = -447.99493
Iteration 5:   log likelihood = -447.99493

Probit regression                                 Number of obs   =       2675
                                                  LR chi2(4)      =     449.31
                                                  Prob > chi2     =     0.0000
Log likelihood = -447.99493                       Pseudo R2       =     0.3340

------------------------------------------------------------------------------
       train |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |  -.0661692   .0062066   -10.66   0.000    -.0783339   -.0540044
        educ |  -.1140225   .0201033    -5.67   0.000    -.1534243   -.0746208
       black |   1.345078   .1150059    11.70   0.000      1.11967    1.570485
        hisp |   .9881963   .2176798     4.54   0.000     .5615517    1.414841
       _cons |   .9944103   .3476386     2.86   0.004     .3130512    1.675769
------------------------------------------------------------------------------

. eststo probit_1

.
. esttab logit_1 probit_1,nogaps p star wide nopar

--------------------------------------------------------------------
                      (1)                             (2)
                    train                           train
--------------------------------------------------------------------
train
age                -0.132***        0.000      -0.0662***        0.000
educ               -0.229***        0.000       -0.114***        0.000
black               2.749***        0.000        1.345***        0.000
hisp                2.177***        0.000        0.988***        0.000
_cons               2.136**         0.002        0.994**         0.004
--------------------------------------------------------------------
N                    2675                            2675
--------------------------------------------------------------------
p-values in second column
* p<0.05, ** p<0.01, *** p<0.001

. esttab logit_1 probit_1 using logit-probit.rtf,nogaps p star wide nopar
(output written to logit-probit.rtf)
```

6.3 Leaving the U.S. Navy – Gruppierte und individuelle Schätzungen LP und LOGIT

Vorhersage des Ausscheidens von Rekruten vor dem Rekrutierungsende (Maddala 1983, S. 32 → Warner 1978)

Modell:

y=1 if the person leaves the Navy before completion of the first enlistment

y=0 otherwise

x years of education (ED), mental ability (MENTAL) (as measured by the Armed Forces Qualification Test, AFQT), martial status, age and race.

Groups: $3 \times 5 \times 3 \times 2 \times 2 = 180$ Zellen aus den Variablenkategorien

n= 30.000 Individuen von 67.000 männlichen Rekruten 1973

Methoden:

Gruppierte Daten, individuelle Daten
Lineares Wahrscheinlichkeitsmodell (LP), LOGIT

Warner, J.T. (1978), The Prediction of Attrition from Military Service, Report CRC345, Center for Naval Analyses, Alexandria, Virginia (USA).

Die Schätzergebnisse des LP- und des LOGIT-Modells bei gruppierten und ungruppierten Daten finden sich in folgender Tabelle

Table 2.2. *A comparison of estimates of parameter values: individual and grouped data*

Variable	Individual linear	Grouped linear	Individual logit	Grouped logit
Ed < 12	−0.105 (17.04)	−0.109 (14.14)	−0.672 (21.23)[a]	−0.656 (14.42)
Ed > 12	0.028 (3.88)	0.032 (3.79)	0.349 (4.51)	0.284 (2.87)
Mental group I	0.084 (9.95)	0.084 (9.65)	1.179 (9.32)	1.040 (6.00)
Mental group II	0.021 (3.96)	0.020 (3.09)	0.201 (4.50)	0.208 (3.60)
Mental group III	−0.053 (7.70)	−0.052 (6.20)	−0.345 (7.71)	−0.342 (6.00)
Mental group IV	−0.098 (12.46)	−0.097 (10.04)	−0.581 (12.98)	−0.571 (9.75)
Dependents	−0.046 (4.82)	−0.039 (3.61)	−0.349 (5.52)	−0.403 (5.21)
Age < 18	−0.031 (4.16)	−0.024 (2.56)	−0.145 (3.24)	−0.166 (3.14)
Age > 19	−0.027 (4.30)	−0.022 (3.51)	−0.185 (4.13)	−0.169 (3.24)
Race (non-Caucasian)	0.027 (3.61)	0.037 (4.15)	0.136 (3.04)	0.081 (1.28)
Constant	0.881 (25.70)	0.882 (20.79)	1.959 (61.96)	1.950 (40.87)
N	30,000	137	30,000	137

[a] *t* values in parentheses.

Quelle: Maddala 1983, S.33

6.4 Nachfrage nach dauerhaften Konsumgütern – Gruppierte Schätzungen LOGIT

Der Einfluss sozioökonomischer Größen auf die individuelle private Nachfrage nach dauerhaften Konsumgütern – Eine Anwendung der diskreten Entscheidungsmodelle LOGIT und TOBIT (Merz 1983)

Modell:

y=1	Kauf eines Konsumgutes j ($j = 1,\ldots,10$)
y=0	sonst
x	soziale Stellung des Haushaltsvorstandes, Alter des HHV, Haushaltsgröße, Bestand an Konsumgütern, Alter des Bestandes, Spar- und Wertpapier vermögen, Haushaltsnettoeinkommen
Groups:	288 Zellen aus n=47.000 Haushaltsrecords der EVS 1969

Methode:

Gruppierte Daten LOGIT / GLS zur Gewichtung unterschiedlicher Zellenbesetzungen

Ergebnisse: Tab. 6.2

Tab. 6.2: Individuelle Nachfrage nach dauerhaften Konsumgütern – Gruppierte Schätzungen LOGIT

Tabelle 2: Die geschätzten Parameter des LOGIT- Ansatzes
zur Erklärung der Ausgabenereignisse für dauerhafte Konsumgüter aller Haushalte in der EVS '69

		SOZIALE STELLUNG DES HAUSHALTSVORSTANDES					ALTER DES HAUSHALTSVORSTANDES		HAUSHALTS-	
DAUERGÜTER KONSUMGUT	1 CONST	2 SSLAND	3 SSSELB	4 SSBEAM	5 SSANGE	6 SSARBE	7 AH3034	8 AH3564	9 HGR2B2	10 HGH3B3
1 KÜHLSCHRANK	-2.35442	-.384648+	-.387519+	-.672868+	.883438+	.844394+	-.158611+	.878711+	.347658	.225614+
2 TIEFKÜHLTRUHE, -SCHRANK	-5.15145	.495599	-.241238+	-.861499+	-.141963+	.188691+	.508397	.316723+	.676744	.791818
3 WASCHMASCHINE, GESCHIRRSPÜLER (ELEKTR.)	-3.84310	.282231+	-.128188+	.886586+	.817532+	.228340+	.189447+	.888631+	.681416	.771988
4 HÖHERWERTIGE HAUSHALTSGERÄTE (ELEKTR.)	-2.37524	-.362986+	-.226492+	.114678+	.154778+	.180581+	-.816882+	-.818467+	.426419	.442566
5 PERSONENKRAFTWAGEN	-4.84310	-.242272+	-.818503+	-.826964+	.841597+	.215442+	1.885219	.424823+	-.613686+	.1189211
6 ZWEIRAD	-3.44430	.823688+	.885375+	.116378+	.112162+	.232048+	.283499+	.230181+	.384514	.953286
7 FERNSEHER	-2.54916	-.178693+	-.543682+	-.818384+	-.189439+	.126738+	.286887+	.885479+	.222291+	.863986+
8 RADIO, KOMBIGERÄT, PLATTENSPIELER	-3.88227	-.412369+	-.721691	.866832+	.125973+	.263832	.342377	.361767	.638176+	.263920+
9 FOTO-, FILMAPPARATE	-3.68926	.888858+	-.267938+	.248588+	.298552+	.358822	.463691	.129819+	.192345+	.162798+
10 IMMOBILIEN	-4.28262	.847736+	.1644621	.365283+	.387689+	.189996+	.594338	.157789+	.476828	.666144

	GRÖSSE	BESTAND VORHANDEN ANSCHAFFUNGSJAHR BEKANNT		SPAR-,WERT-PAPIER-VERMÖGEN (T-1)	HAUSHALTS-NETTOEINKOMMEN
DAUERGÜTER KONSUMGUT	11 HGB41M	12 ABEXIS	13 ADHEKA	14 SPWERE	15 HHTINK
1 KÜHLSCHRANK	.16359+	-.288862+	-1.166432	.888159+	.888833
2 TIEFKÜHLTRUHE, -SCHRANK	.93882	2.346895+	-2.619698+	-.888169+	.888842+
3 WASCHMASCHINE, GESCHIRRSPÜLER (ELEKTR.)	.98487	-.423394	-4.795148+	-.888209+	.888821+
4 HÖHERWERTIGE HAUSHALTSGERÄTE (ELEKTR.)	.46854	-.188428+	.156288+	-.888895+	.888826+
5 PERSONENKRAFTWAGEN	-.82573+	.666624+	-.145178+	.888184+	.888872
6 ZWEIRAD	1.57528	.573784	.224812+	-.888668+	.888887+
7 FERNSEHER	.88698+	-.584898	.389586+	-.888146+	.888839
8 RADIO, KOMBIGERÄT, PLATTENSPIELER	.18387+	.874778+	-.478878+	-.888828+	.888862
9 FOTO-, FILMAPPARATE	.67182+	.643817	.532906	-.888843+	.888845
10 IMMOBILIEN	.73987	1.288898	-.849511+	-.888881+	.888821+

Irrtumswahrscheinlichkeiten
t-Wert < 0.20, α 0.1 %
* t-Wert < 2.57, α 1.0 %
* t-Wert < 1.96, α 5.0 %
\+ t-Wert ≥ 1.96, α > 5.0 %

Quelle: Merz 1983, S. 240, Eigene Berechnungen

Fortsetzung Tab. 6.2

Endogene Variablen:

1. Kühlschrank
2. Tiefkühltruhe, -schrank
3. Elektrische Waschmaschine (mit und ohne Schleuder), Wäscheschleuder, Geschirrspülmaschine
4. Elektrische Nähmaschine, Staubsauger (incl. Elektrischer Teppichklopfer), elektrischer Grill, elektrischer Heimbügler
5. Personenkraftwagen
6. Zweirad (incl. Motorrad, -roller, Moped, Motorfahrrad (Mofa), Fahrrad)
7. Fernsehgerät (auch Fernsehtruhe, Phonokombination (Truhe) mit Fernsehgerät)
8. Rundfunkgerät (auch Musiktruhe), Tonbandgerät, Plattenspieler, Phonokombination (Truhe) ohne Fernsehgerät)
9. Foto-, Filmapparat, Projektor für Dia oder Film
10. Grundstück, Gebäude, Eigentumswohnung u. ä.

Exogene Variablen

1. Soziale Stellung des Haushaltsvorstandes

Landwirt	(SSLAND)
Selbständiger	(SSSELB)
Beamter	(SSBEAM)
Angestellter	(SSANGE)
Arbeiter	(SSARBE)
Nichterwerbstätiger	(SSNERW)

2. Alter des Haushaltsvorstandes

0 bis 34 Jahre	(AH0034)
35 bis 64 Jahre	(AH3564)
65 Jahre und älter	(AH65AL)

3. Haushaltsgröße

Einpersonenhaushalt	(HG0101)
Haushalt mit 2 Personen	(HG0202)
Haushalt mit 3 Personen	(HG0303)
Haushalt mit 4 und mehr Personen	(HG04UM)

4. Jeweiliger Bestand der Gebrauchsgütergruppe zum Ende der Vorperiode

Kein Bestand vorhanden	(ABNEXI)
Bestand vorhanden	(ABEXI)

für Immobilien: Wohnstatus

Mieter der Wohnung

Eigentümer der Wohnung

5. Anschaffungsjahr des jeweiligen Bestandes

vor 1958 Anschaffungsjahr nicht bekannt	(AJNIBE)
Anschaffungsjahr bekannt	(AJBEKA)

für Immobilien: Baujahr des Gebäudes

1958 und später

vor 1958

Quantitative Variablen:

6. Spar- und Wertpapiervermögen, sowie der Wert der Geschäfts- und Genossenschaftsanteile zum Ende der Vorperiode (SPWEGE)
7. Haushaltsnettoeinkommen (HHEINK)

6.5 Travelling by Auto or Transit – Individualschätzung LOGIT

Reisealternativen/Transportwesen: Wichtiges Anwendungsgebiet diskreter Entscheidungsmodelle.

Domencich, T. und D. McFadden 1975, Urban Travel Demand – A Behavioral Analysis, Amsterdam; Metropolitan Washington area (Cambridge Sytematics 1976, Ben-Akiva und Lerman 1991, S. 77 ff)

Modell:

$y=1$ Auto
$y=0$ Transit
x In-vehicle and out-of vehicle time, costs, transit fare, auto ownership, downtown workplace

Daten:

n=1476 worktrips, Metropolitan Washington area

Methode:

Individualschätzung LOGIT, Maximum Likelihood

Ergebnisse:

Text und Tab. 6.3

Tab. 6.3: Travelling by auto or transit

Binary Logit Example

Our second example is a binary logit model using data from a 1968 survey in the Metropolitan Washington area (Cambridge Systematics 1976). (This large data base will be used for a subsequent example in this book—a multiple-choice model.) As in Lisco's model the two travel alternatives are auto and transit. In this case, however, transit is the bus mode rather than rail rapid transit.

The specification of the model is summarized in table 4.2. Note that here there are seven coefficients labeled $\beta_1, \ldots, \beta_7$. Rather than discuss each variable in detail, we consider only the relevant differences between the utilities defined in tables 4.1 and 4.2.

1. Travel time has been divided into two components, in-vehicle and out-of-vehicle time. This division reflects the observation that travelers would generally prefer to spend a minute traveling in a vehicle than a minute waiting for or walking to one. In the case of the transit mode, out-of-vehicle time includes walking to a bus stop, waiting for the vehicle, any transfer time, and walking from a bus stop to the final destination. Auto out-of-vehicle time includes time spent going to and from the vehicle.

The coefficients of both time variables are negative, as expected, and the coefficient of out-of-vehicle time is about two and a half times that of in-vehicle time. This ratio is typical for work mode choice models.

2. The two time-related variables are treated as generic; however, the cost variable is treated as *alternative specific*. Basically it is assumed that a minute has the same marginal effect on each alternative's utility, whereas a cent in fare and out-of-pocket cost have *different* marginal effects. This may be attributable to differences in how people perceive auto costs and transit fares, since the former are not usually incurred directly out of pocket with each trip made.

3. The alternative-specific socioeconomic variables used are entirely different from those used by Lisco. Here household automobile ownership and a dummy variable for whether the traveler works downtown are used. The auto ownership variable obviously reflects the important effect of having a large number of autos available in choosing to use the auto mode. The dummy variable for whether or not the traveler works downtown reflects the fact that auto use to downtown areas has associated with it a high degree of travel time unreliability. Even though the mean time may be identical for a suburban- and urban-destined worker, the variance of the urban traveler's time will generally be much greater. The uncertainty about arrival time may induce downtown workers to leave from home somewhat earlier than their suburban counterparts in order to avoid being late for work too frequently. This extra time may be viewed as a source of disutility to the urban traveler (see Abkowitz 1980).

Table 4.2
Binary logit model: $P_n(A) = e^{V_{An}}/(e^{V_{An}} + e^{V_{Tn}})$

	β_1	β_2	β_3	β_4	β_5	β_6	β_7
Auto utility, V_{An}	1	Auto in-vehicle time (min)	Auto out-of-vehicle time (min)	Auto out-of-pocket cost (¢)	0	Household auto ownership	1 if workplace is in downtown Washington; 0 otherwise
Transit utility, V_{Tn}	0	Transit in-vehicle time (min)	Transit out-of-vehicle time (min)	0	Transit fare (¢)	0	0
Estimated coefficients	1.454	−0.00897	−0.0308	−0.0115	−0.00708	0.770	−0.561

Quelle: Ben Akiva, 1983

6.6 Arbeitsangebot von Frauen – PROBIT

Female Labor Supply - Labor Force Participation, Market Wage Rate and Working Hours of Married and Unmarried Women in the Federal Republic of Germany (Merz 1990).

Selektionskorrigierte Schätzung (à la Heckman) des Arbeitszeitangebots (vgl. hier Kap. III) über die Partizipationsentscheidung.

Modell:

$y=1$ Frau geht arbeiten
$y=0$ sonst
x sozioökonomische Charakteristika

- der Frau selbst
- des Ehemanns
- der Haushaltssituation

→ own wage and income elasticities – Berechnungen

Daten:

Sozio-ökonomisches Panel, Welle 1 (1984):

$n_{\text{Personen}} = 12.246$ $n_{\text{Haushalte}} = 5.921$ $n_{\text{married women}} = 3.932$ $n_{\text{unmarried women}} = 2.176$

Methode:

Individualschätzung PROBIT, Maximum Likelihood

Ergebnisse:

Tab. 6.6a Arbeitsangebot von Frauen – PROBIT
Tab. 6.6b Arbeitsangebot von Frauen - Elastizitäten

Tab. 6.6a: Arbeitsangebot von Frauen - PROBIT

Table 1: Labor force participation of married and unmarried women in the Federal Republic of Germany (FRG): Estimated coefficients of the PROBIT-Indexfunction

	Participation			
	married		unmarried	
	coeff.	t-value	coeff.	t-value
PERSONAL CHARACTERISTICS				
Family status				
separate (married)	-	-	-0.122	(0.54)
divorced	-	-	-0.171	(1.22)
widowed	-	-	-0.154	(0.89)
Age	-0.020	(0.28)	0.210	(3.71)**
Age $^2 \cdot 10^{-2}$	-0.027	(0.16)	-0.510	(4.01)**
Age $^3 \cdot 10^{-4}$	-0.048	(0.37)	0.283	(3.20)**
Work experience				
Full-time years	0.095	(11.41)**	0.081	(5.46)**
Years ^{2}	-0.067	(2.86)**	-0.078	(2.35)*
Interruptions	0.134	(2.34)**	0.081	(0.91)
Max. interruptions	0.022	(2.59)**	0.016	(1.23)
Part-time years	0.153	(14.45)**	0.118	(5.16)**
Years ^{2}	-0.284	(6.47)**	-0.182	(2.26)*
Interruptions	0.145	(1.31)	-0.199	(0.88)
Max. interruptions	-0.029	(1.70)*	0.030	(1.40)
Vocational and further training	-0.260	(1.42)	-0.531	(5.12)**
School-leaving certificate (diploma)				
Intermediate (Realschule)	0.201	(2.72)**	0.426	(4.79)**
Specialized upper secondary (Fachhochschulreife)	0.087	(0.42)	-0.071	(0.35)
Upper (Abitur)	0.124	(0.72)	0.117	(0.83)
Completed vocational training				
Apprenticeship (Lehre)	-0.084	(1.35)	0.147	(1.61)
Specialized vocational (Berufsfachschule)	0.003	(0.04)	0.156	(1.24)
Specialized college of higher education (Fachhochschule)	0.286	(1.26)	-0.133	(0.33)
University	0.309	(1.42)	0.607	(2.76)**
Civil servant training	0.272	(0.93)	0.247	(0.68)
Health service	0.293	(1.92)*	0.105	(0.45)
Foreigner	0.256	(3.79)**	-0.043	(0.42)
Health situation				
Impairment [1]	0.039	(0.79)	0.036	(0.45)
Chronic Complaints [2]	-0.110	(1.56)	-0.097	(0.91)
Certified handicapped [2]	-0.106	(0.84)	-0.197	(1.15)
(Subjective) satisfaction [3]				
Household income	0.014	(1.29)	0.041	(3.00)**
Household activities	-0.033	(2.83)**	-0.014	(0.82)
Monthly unearned income:				
Old age, widow's pension [6]	-1.638	(5.39)**	-0.777	(6.10)**
Private support [4][6]	-0.222	(0.21)	-0.498	(2.23)*
Unemployment benefit/relief [6](t-1)	-0.652	(3.83)**	-1.075	(5.09)**
HUSBAND'S CHARACTERISTICS				
Occupational status				
Manual worker	-0.176	(1.81)*	-	-
White-collar worker	-0.320	(2.86)**	-	-
Civil servant	-0.357	(2.63)**	-	-
Self-employed	-0.235	(1.94)*	-	-
Net earned income last month [6]	0.091	(3.80)**	-	-

Fortsetzung Tab. 6.6a: Arbeitsangebot von Frauen - PROBIT

	Participation			
	married		unmarried	
	coeff.	t-value	coeff.	t-value
HOUSEHOLD CHARACTERISTICS				
Nursing cases existing [2]	0.014	(0.08)	0.154	(0.76)
Children in age classes				
0 to 5 years	-0.595	(11.44)**	-0.259	(1.76)*
6 to 9 years	-0.205	(3.48)**	-0.163	(1.25)
10 to 15 years	-0.103	(2.48)**	0.102	(1.46)
16 and older	-0.018	(0.54)	-0.002	(0.05)
Community size (Boustedt)				
up to 20.000 Einwohner	0.022	(0.29)	0.077	(0.67)
100.000 inhabitants and more	0.008	(0.12)	-0.097	(0.98)
Monthly unearned income: transfer payments				
Housing benefits [6]	-1.171	(1.28)	-2.621	(2.42)**
Welfare benefits [5][6]	-0.996	(1.46)	-1.336	(2.92)**
Yearly unearned income: property				
Income from let and lease [6]	-0.018	(2.74)**	-0.016	(1.93)*
Income from interests and dividends [6]	-0.008	(0.72)	0.060	(1.80)*
Constant	0.493	(0.53)	-2.503	(3.41)**
Pseudo R^2	0.300	-	0.400	-

level of significance: * (95%), ** (99%)

1) 0 = not at all, 1 = slightly, 2 = considerably

2) 0 = no, 1 = yes

3) scale 0-10, 0 = completely discontent, 10 = completely content

4) from private persons not living in the household

5) inclusive subsistence benefits/benefits granted in needy situations

6) $*10^{-3}$

Source: Socio-Economic Panel 1984, Wave 1, own computations

Quelle: Merz 1990

Tab. 6.6b: Arbeitsangebot von Frauen – Elastizitäten

Table 3: Own wage (uncompensated and compensated), total income and (property) income elasticities of labor force participation and working hours for married and unmarried women in the Federal Republic of Germany (FRG)

	Own wage and income elasticities			
	Participation Probability PROBIT		Working hours PROBIT/COLS	
Independent/Depeadent variables	married	unmarried	married	unmarried
PERSONAL CHARACTERISTICS				
Own wage[1)]				
uncompensated (a)	3.30254**	1.63248**	-.10147*	-.07723**
compensated (b)	3.30193	1.63241	-.08172	-.07505
Total income[1)] (c)	.00061	.00007	-.01975	-.00218
Personal transfers				
Old age, widow's pension	-.847**	-.925**	.048	-.279**
Private support	.447	-.426**	-.275	-.191**
Unemployment benefits (t-1)	.018	-.475**	.004	-.010
HUSBAND CHARACTERISTICS				
Net earned income last month	.588**	-	-.141**	-
HOUSEHOLD CHARACTERISTICS				
Household Transfers				
Housing benefits	.008	-.195	-.010	-.104*
Welfare benefits	-.148	-.378*	.160	.080
Yearly unearned income: property				
Income from let and lease	-.449**	-.302**	-.109*	.061*
Income from interests and dividends	-.033*	.047	.006	-.012

Participation (probability) elasticities:
$\epsilon_{prob,xi} = \phi(X\beta)\beta_i \, x/p$ with $p = \Phi(X\beta)$ = working/all

Hours elasticities:
$\epsilon_{H/xi} = \beta_i \, x/H$

Elasticities with estimated ln wages:
$\epsilon_{prob,w} = \phi(X\beta)\,\beta_{lnw}\,1/p$

$\epsilon_{H,w} = \beta_{lnw}\,1/H$

1) The labor supply (H_i) elasticities are computed out of
$(\partial H_i/\partial W_i)\,W/H = S_{wi}(H_i)\,W/H + (\partial H_i/\partial V)\,W$, that is $\beta_{lnw}\,1/H = S_{wi}(H_i)\,W/H + W\sum_j \beta_{vj}$
(a) (b) (c)
where
(a) uncompensated (gross) wage elasticity
(b) income compensated (substitution) elasticity
(c) total income elasticity.

The appropriate participation elasticities are computed analogously out of
$(\partial p_i/\partial W_i)\,W/p = S_{wi}\,(p_i)\,(W/p) + (\partial p_i/\partial V)\,W$, that is $\phi\,(X\beta)\,\beta_{lnw}\,1/p = S_{wi}\,(p_i)\,W/p + \phi\,(X\beta)\,W\sum_j \beta_{vj}$.

with j as unearned (property) income components here as:
husband's income (if married), income from let and lease, income from interests and dividends.

All elasticities are calculated at sample mean.

Source: Socio-Economic Panel 1984, Wave 1, own computations

Quelle: Merz 1990

III MULTIPLE ENTSCHEIDUNGSMODELLE

1 Struktur multipler Entscheidungsmodelle

Bisher: binäre Entscheidung für eine Kategorie, eine Auswahl (z.B. Busfahren, ja oder nein)
Nun: multiple Kategorien, Set von Wahlentscheidungen, Alternativen
$j = 1,\ldots,m$ z.B. Bus, Zug, Auto oder … fahren

Beispiele:

- Transportmittelwahl (Bus, Bahn, Auto, …)
  Domencich und McFadden 1975, Theil 1969
- Portfoliostruktur (Vermögensanlagen)
  Uhler und Cragg 1971
- Berufswahl (menial („niedere Dienste"), blue collar, craft, white collar, professional)
  Schmidt und Strauss 1975
- Zuordnung von bond ratings als eine Auswahl von multiplen Alternativen
  Terza 1985
- …

Es steht also bei multiplen Entscheidungsmodellen die Wahl *einer* Alternative *j* aus einem Set von *j* = 1, …, *m* Alternativen im Vordergrund.

2 Nutzentheoretische Fundierung: RUM – Random Utility Maximization

RUM: Random Utility Maximization

RUM-Modell (McFadden 1976, 1985, Domencich und McFadden 1975; Survey Amemiya 1981):
Allgemeiner Ansatz des stochastisches Nutzenmodells bei m diskreten Alternativen:
Rationaler Entscheider wählt eine diskrete Alternative j über die Maximierung seines Nutzens.

- Jedes der *n* Individuen *i* ($i = 1,\ldots,n$) hat die Wahl zwischen *m* diskreten Alternativen *j* ($j = 1,\ldots,m$)
- Jede Alternative hat einen Nutzenindex

  $u_{ij} = v_{ij} + \varepsilon_{ij}$

  $u_{ij} = \alpha' x_j + \beta_j' w_i + \gamma' z_{ij} + \varepsilon_{ij}$ $\quad v_{ij}$ deterministisch, ε_{ij} zufällig

  $u_{ij} = \beta_{(j)}' x_i + \varepsilon_{ij}$
- Deterministischer Nutzen

  $v_{ij} = \alpha' x_j + \beta_j' w_i + \gamma' z_{ij}$

mit

x_j = Vektor der Eigenschaften der Alternative j (z. B. Preise, Geschwindigkeit bei Alternativen Bus, Bahn, Auto)

w_i = Vektor sozioökonomischer Größen des Individuums i (z. B. Alter, Geschlecht, Einkommen)

z_{ij} = Vektor der Größen, die mit den Individuen i und Alternativen j variieren

α, β_j, γ = zu schätzende Parametervektoren

- Zufallsterm

  ε_{ij} = f (unbeobachtete Charakteristika der Alternativen j bzw. der Individuen i, Messfehler, Verwendung von Instrumentvariablen für unbeobachtete Größen)

- Entscheider wählt die Alternative *j* aus, die ihm den höchsten Nutzen stiftet:

  $u_{ij} > u_{ik}\ (j,k = 1,...,m;\ j \neq k)$

  bzw.

  $u_{ij} = \max(u_{i1},\ldots,u_{im})$

  oder nach Umformung

  $v_{ij} + \varepsilon_{ij} > v_{ik} + \varepsilon_{ik}$

  $\left(v_{ij} - v_{ik}\right) > \left(\varepsilon_{ik} - \varepsilon_{ij}\right) \qquad (für\ j \neq k)$

- Auswahlwahrscheinlichkeit:

  $P_{ij} = \text{Prob}\left[\varepsilon_{ik} - \varepsilon_{ij} < v_{ij} - v_{ik}\right] \qquad (j,k = 1,\ldots,m;\ j \neq k)$

  Differenzen in Bezug auf eine Basisalternative k.

RUM-Model für das Multinomiale Logit

$$u(Wahl\ von\ j\ für\ Individuum\ i) = u_{ij} = \beta' x_{ij} + \varepsilon_{ij} \qquad (j = 1,...,m)$$

Die Fehlerterme $(\varepsilon_{i1}, \varepsilon_{i2}, ..., \varepsilon_{im})$ sind annahmegemäß unabhängig und jeder Fehlerterm mit einer Extremwert (Gumbel)-Verteilung verteilt. Die Wahrscheinlichkeit dafür, dass das Individuum i die Alternative j wählt ist

$$\text{Prob}(u_{ij} > u_{ik}) \quad für\ alle\ k \neq j\ .$$

Es kann gezeigt werden, dass für die unabhängige Extremwertverteilung diese Wahrscheinlichkeit

$$\text{Prob}(y_i = j) = \exp(\beta' x_{ij}) / \sum_j \exp(\beta' x_{ij})$$

ist.

RUM und das Modell der Indexfunktionen bei m Alternativen

Nicht beobachtete Zufallsvariable $Z = (u_1, ...u_m)'$

Max. Nutzenindex u_j *bzw.* v_j aus $v_K, (k = 1, ...m)$

wählt Ereignis j aus:

$$y_j = \begin{cases} 1 & wenn \quad u_j - u_K > 0 \qquad j,k = 1,...m \quad j \neq k \\ 0 & sonst \end{cases}$$

$$\text{Prob}(y_j = 1) = \text{Prob}(Z \in \text{best. Nutzendifferenzen})$$

Die Nutzendifferenz ist eine kontinuierliche Zufallsvariable, die, wenn sie über einen Schwellenwert hinausgeht, über die Indexfunktion den Eintritt eines Ereignisses *j* bestimmt. Die dahinter liegenden Nutzenkomponenten sind nicht direkt beobachtbar.

3 Das MULTINOMIALE LOGIT (MNL-)Modell

Im MUTINOMIALEN LOGIT Modell werden typischerweise individuenspezifische Merkmale x_i zur Erklärung herangezogen.

Seien $P_1, P_2, \ldots, P_m$ die Wahrscheinlichkeiten für die *m* zur Auswahl stehenden Alternativen. Diese Wahrscheinlichkeiten können in „binärer" Form als Verhältnis zweier Wahrscheinlichkeiten geschrieben werden (Maddala 1983, S. 34 ff):

$$\frac{P_1}{P_1 + P_m} = F(x_i\beta_1) \; ; \; \frac{P_2}{P_2 + P_m} = F(x_i\beta_2) \; ; \; \frac{P_{m-1}}{P_{m-1} + P_m} = F(x_i\beta_{m-1})$$

Nach Umformung über

$$\frac{P_j}{P_j + P_m} = F(x_i\beta_j) \; ; \; \frac{P_j}{F(x_i\beta_j)} = P_j + P_m \; ; \; \frac{1}{F(x_i\beta_j)} = 1 + \frac{P_m}{P_1}$$

$$\frac{P_m}{P_1} = \frac{1}{F(x_i\beta_j)} - 1; \; \frac{P_m}{P_j} = \frac{1 - F(x_i\beta_j)}{F(x_i\beta_j)}$$

ergibt sich

$$\frac{P_j}{P_m} = \frac{F(x_i\beta_j)}{1 - F(x_i\beta_j)} = G(x_i\beta_j) \quad (j = 1, \ldots, m-1).$$

Wegen $\sum_{j=1}^{m-1} \frac{P_j}{P_m} = \frac{1 - P_m}{P_m} = \frac{1}{P_m} - 1$

ist $P_m = \frac{1}{1 + \sum_{j=1}^{m-1} G(x_i\beta_j)}$

und damit die Wahrscheinlichkeit für die Alternative j im MULTINOMIALEN MODELL

$$P_j = \frac{G\left(x_i\beta_j\right)}{1+\sum_{k=1}^{m-1} G\left(x_i\beta_k\right)}$$

Die Beobachtungen stammen aus einer multinominalen Verteilung mit den Wahrscheinlichkeiten P_j und P_m.

In Abhängigkeit wieder von der Wahl von $G\left(x_i\beta_j\right)$ ergibt sich das

MULTINOMIALE LOGIT (MNL)

$$G\left(x_i\beta_j\right) = \mathrm{e}^{x_i\beta_j}$$

MNL: $P_j = \frac{e^{x_i\beta_j}}{1+\sum_{k=1}^{m-1} e^{x_i\beta_k}}$ oder das

MULTINOMIALE PROBIT (MNP)

$G\left(x_i\beta_j\right)$ bzw. P_j als mehrdimensionales Integral

Die Schwierigkeit das MNP zu schätzen liegt an den multivariat normal verteilten Wahrscheinlichkeiten, die mit einer Dimensionalität von größer 2 nur sehr aufwendig mit neueren Ansätzen von ‚simulated estimators' machbar ist. Vertiefungen und Hinweise finden sich dazu bspw. in Greene 2003, S. 727 ff oder für den MNP Ansatz generell Maddala 1983, S. 62 ff

Ein unterschiedlicher Bezug der β auf Charakteristika der Alternativen und/oder individuelle Merkmale führt zu verschiedenen ‚MNL'-Ansätzen:

$x_{ij}\beta$: β „über alles", weder individuen- noch alternativenspezifisch

(Boskin 1974: several occupations, jeweils charakterisiert durch 3 Variablen
– present value of potential earnings
– training cost/net worth
– present value of time unemployed)

$x_i\beta_j$: β alternativenspezifisch/„normales MNL"
(Schmidt und Strauss 1975a, m occupations y_i, x_i = age, sex, years of schooling etc.)

$x_{ij}\beta + w_i\alpha_j$: generelles Modell aus beiden Ansätzen

(McFadden 1974: shopping mode choices
x_{ij}: alternativenspezifisch/individuenspezifisch: transit costs, transit walk time

(individuelle Entfernung zu den Haltestellen unterschiedlicher Transportmittel)…

w_i : individuenspezifisch: race, occupation, …)

Je nach Spezifikation sind die geschätzten Koeffizienten nicht immer leicht zu interpretieren. Stehen wie im Ausgangsmodell die $x_i\beta_j$ im Erkenntnisinteresse, dann gehen in die marginalen Wahrscheinlichkeiten aus β_j nicht nur die Alternative j ein, sondern auch alle anderen Alternativen (siehe unten: marginale Effekte).

4 McFadden's CONDITIONAL LOGIT

Beziehen sich die Daten auf alternativenspezifische Attribute, dann ist das angemessene Modell multipler Entscheidungsmöglichkeiten das CONDITIONAL LOGIT Modell von McFadden. Die nutzentheoretische Fundierung ist das RUM-Modell mit: $y_j = 1 \text{ wenn } \mathrm{u_j} = y_j^* = \max\left(y_1^*,\ldots,y_m^*\right)$, sonst ist $y_j = 0$.

Der stochastische Nutzen ist $u_{ij} = y_{ij}^* = \upsilon_{ij}\left(z_{ij}\right) + \varepsilon_{ij} = z_{ij}\beta + \varepsilon_{ij}$ .Der Nutzen ist somit eine Funktion von z_{ij}, einem Vektor, der sowohl individuenspezifische als auch alternativenspezifische Charakteristika enthält.

McFadden 1973 hat gezeigt, dass nur bei einer **Type I Extremwertverteilung (Gumbel Verteilung)** $f(\varepsilon) = \exp\left[-\varepsilon - \exp(-\varepsilon)\right]$ der unabhängigen Fehler ε mit $F\left(\varepsilon_j < \varepsilon\right) = \exp\left[-\exp(-\varepsilon)\right]$ das CONDITIONAL LOGIT Modell folgt

CONDITIONAL LOGIT

$$P\left(y_j = 1 \middle| z_{ij}\right) = \frac{e^{\upsilon_{ij}}}{\sum_{j=1}^{m} e^{\upsilon_{ij}}} \triangleq \frac{\exp\left(z_{ij}\beta\right)}{1 + \sum_{k=1}^{m-1} \exp\left(z_{ij}\beta\right)}$$

Sei $z_{ij} = (x_{ij}, w_i)$, dann variiert x_{ij} über die Alternativen (und möglicherweise auch über die Individuen); typischerweise Attribute der Alternativen. w_i enthält die Charakteristika der Individuen. Dann wird McFadden's CONDITIONAL LOGIT Modell zu

$$y_{ij}^* = u_{ij} = \underbrace{x_{ij}\beta + w_i\alpha_j}_{\upsilon_{ij}} + \varepsilon_{ij}$$

und

$$P_j = \frac{\exp\left(x_{ij}\beta + w_i\alpha_j\right)}{\sum_{k=1}^{m} \exp\left(x_{ik}\beta + w_i\alpha_k\right)}.$$

Durch geeignete Multiplikation der Individuenspezifika mit Dummyvariablen können alternativenspezifische individuelle Charakteristika berücksichtigt werden. Eine der Interaktionsterme ist wegen der linearen Abhängigkeit auszulassen.

Beispiel: Wahl eines Einkaufszentrums aus mehreren Alternativen.

Die Attribute der Einkaufszentren (Anzahl der Geschäfte, Entfernung zur City) sind für alle Individuen gleich, nicht aber ein Einkommen, das indivduenspezifisch ist. Mit drei Einkaufszentren ließe sich ein CONDITIONAL LOGIT Modell formulieren mit

Alternative j	alternativenspezifisch		individuenspezifisch	
Zentrum 1:	Anzahl der Geschäfte	Entfernung	Einkommen	0
Zentrum 2:	Anzahl der Geschäfte	Entfernung	0	Einkommen
Zentrum 3:	Anzahl der Geschäfte	Entfernung	0	0

Die Erhebung alternativen- *und* individuenspezifischer Charakteristika (z.B. Entfernung des Individuums zum jeweiligen Zentrum) ist aufwendig; entsprechende Modelle sind daher eher selten anzutreffen.

5 Schätzung des MULTINOMIAL LOGIT und des CONDITIONAL LOGIT Modells

Die Schätzung der Koeffizienten $\beta_{(j)}$ des MULTINOMIALEN LOGIT und des CONDITIONAL LOGIT Modells wird wieder über den Maximum-Likelihood Ansatz vorgenommen.

Sind m-1 Wahrscheinlichkeiten bekannt, dann ist die verbleibende Wahrscheinlichkeit P_m leicht aus $1-\sum_{i=1}^{m-1} P_i$ zu berechnen. Deshalb ist zur eindeutigen Identifikation eine Normierung notwendig (z.B. $\beta_m = 0$ oder $\beta_j^* = \beta_j - \beta_m$)

Likelihood Funktion

$$L = \prod_{i=1}^{n} P_{i1}^{y_{i1}} P_{i2}^{y_{i2}} \dots P_{im}^{y_{im}}$$

$$\ln L = \sum_i \sum_j y_{ij} \ln \text{Prob}\left(y_i = j\right) \qquad | \triangleq \text{LOGIT (MNL), PROBIT (MNP)}$$

1. Ableitungen

$$\frac{\partial \ln L}{\partial \beta_j} = \sum_i \left(y_{ij} - P_{ij}\right) x_i = 0$$

wobei $y_{ij} = 1$ wenn von Individuum i die Alternative j gewählt wurde, 0 sonst.

Nichtlineares Gleichungssystem z.B. zu lösen über Newton $\rightarrow \hat{\beta}_j$

Die Hesse'sche Matrix der 2. Ableitungen hat

$(m-1)^2$ Blöcke $(k+1)(k+1)$

j-ter Diagonal Block: $\sum_i -P_{ij}\left(1-P_{ij}\right)x_i x_i'$

jk Nebendiagonal Block: $\sum_i \left(P_{ij}P_{ik}\right)x_i x_i'$ (Greene 1993, S. 667)

ist negativ definit $\Rightarrow$ Newton-Iterationen konvergieren zu einem eindeutigen Maximum.
Bei der Spezifikation des CONDITIONAL LOGIT Modells ist darauf zu achten, dass eine Konstante nicht vorkommen darf.

Marginale Effekte, Elastizitäten

Der Effekt einer Änderung der unabhängigen Variablen auf die Auswahlwahrscheinlichkeit kann mit den **marginalen Effekten** ermittelt werden. Da es sich wie in den vorangegangenen Modellen um ein nicht-lineares Modell handelt, sind die marginalen Effekte auch hier nicht konstant, d.h. sie sind abhängig von allen anderen Variablenwerten.

Marginale Effekte: $\frac{\partial P_j}{\partial x_i} = P_j\left[\beta_{ij} - \sum_k P_k \beta_{ik}\right]$

Da die marginalen Effekte auch durch die Koeffizienten aller anderen Alternativen beeinflusst werden, kann mit einem Koeffizienten nicht definitiv der direkte Einfluss auf die Wahrscheinlichkeit und die Richtung über das Vorzeichen bestimmt werden.

Da Dummies keine infinitesimalen Änderungen haben, würde man marginale Effekte von Dummy-Variablen berechnen aus der Differenz:

Marginaler Effekt bei Dummies: $\mathrm{P}(y=1 \mid x_{dummy}=1) - \mathrm{P}(y=1 \mid x_{dummy}=0)$.

Die prozentualen Effekte auf die Auswahlwahrscheinlichkeit erfolgen über **Elastizitäten.** Sie werden berechnet durch die Multiplikation des marginalen Effekts mit dem Wert der Regressorvariablen dividiert durch die Wahrscheinlichkeit P_j.

Das **relative Risiko** Alternative j anstatt Alternative 1 zu wählen ist

$$\frac{\Pr(y_i = j)}{\Pr(y_i = 1)} = e^{x_i \beta_j}.$$

Damit gibt ähnlich dem binären Entscheidungsmodell e^{β_j} die proportionale Änderung dieses relativen Risikos an wenn x_j sich um eine Einheit sich ändert.

Schätzgüte

- t-Tests (Koeffizienten)

- LR-Test (Gesamtmodell)
- Predictions – Actual Numbers
  … siehe Kapitel II.5

Choice Based Sampling

Anstatt alle Beobachtungen insgesamt zufällig auszuwählen werden bei „choice based sampling" die Beobachtungen auf der Basis der diskreten Alternative (binär oder multinomial) jeweils (zufällig) ausgewählt. Damit können bspw. relative wenige Beobachtungen für eine Alternative „oversampled" werden, um bessere (effizientere Schätzungen der Parameter) Resultate zu erhalten.

Ist allerdings eine Selektion endogen, dann sind OLS Ergebnisse verzerrt und inkonsistent (Wooldridge 2009, 223). Lösung: Gewichtete (OLS-)Schätzung oder ML Schätzung der gemeinsamen Dichte der unbekannten Parameter und der Stichprobengewichte (Cameron und Trivedi 2005, Kapitel 24.4 Endogenous Stratification).

Stichworte: Endogenous stratification; sampling scheme with endogenous components; sample selection based on dependent variables

6 Unabhängigkeit von irrelevanten Alternativen: IIA-Annahme

Eine Eigenschaft des MNL-Modells (auch des Conditional Logit Modells) ist die, dass die sog. „odds"

$$\frac{P_j}{P_k} = \frac{F(x_i\beta_i)}{1 - F(x_i\beta_i)}$$

unabhängig von irrelevanten Alternativen sind. D.h., die Relation zwischen den Wahrscheinlichkeiten sind unabhängig von der Gesamtanzahl aller Wahlmöglichkeiten *m*. Würde *m* erweitert, verändert das die odds nicht.

Würde also bspw. ein neues Verkehrsmittel auftreten, so könnte man – ohne Neuschätzung – mit den entsprechenden Charakteristika die neue P_{j+} sofort berechnen.

IIA-Ursache: Unabhängigkeits-Annahme der ε_j

IIA-Test: Hausman und McFadden 1984

$$\chi^2 = \left(\hat{\beta}_s - \hat{\beta}_f\right)' \left[\hat{V}_s - \hat{V}_f\right]^{-1} \left(\hat{\beta}_s - \hat{\beta}_f\right) \quad \Big| \, \text{Chi}^2 - \text{verteilt}$$

s steht für eingeschränkte (restricted) Menge der Alternativen, f für vollständiges Set; $\hat{V}_s, \hat{V}_f$ sind die asymptotischen Kovarianzmatrizen von $\hat{\beta}_s, \hat{\beta}_f$. Eine Differenz $\hat{\beta}_s - \hat{\beta}_f = 0$ ist die Nullhypothese, also die für den Test berechenbare Testgröße. Erbringt eine Schätzung des vollen Modells (full model f) keinen Unterschied zur Schätzung des eingeschränkten Modells (bei Auslassen einer Alternative), dann ist die IIA-Annahme erfüllt; damit ist das Multinomiale Modell das pas-

sende Schätzmodell. Die Anzahl der Freiheitsgrade ist gleich der Anzahl der Koeffizienten in $\hat{\beta}_s \, bzw. \hat{\beta}_f$.

Für die konkrete Durchführung des IIA-Tests würde man in einem Fall mit drei Alternativen als vollständiges Set alle drei Alternativen als Ausgangsmodell verwenden. Für das eingeschränkte Set würde man eine Alternative streichen, d.h. das Modell mit zwei Alternativen schätzen.

Aus den beiden Modellschätzungen erhält man jeweils einen Koeffizientenvektor. Nun werden die geschätzten Koeffizienten verglichen:

- β_f Koeffizienten der Schätzung mit vollem Alternativenset (jedoch nur die Koeffizienten für Alternative 1 und 2)
- β_s Koeffizienten der Schätzung mit beschränktem Alternativenset (es liegen nur die Koeffizienten für Alternative 1 und 2 vor)

Je nach Ergebnis ist Alternative 3 irrelevant oder nicht (irrelevant in Bezug auf das Chancenverhältnis der Alternativen 1 und 2). Im Prinzip müsste jede andere Konstellation dann auch noch getestet werden.

Wäre die IIA-Annahme nicht erfüllt, so wäre als Alternative ein MULTIVARIATES LOGIT/PROBIT-Modell mit

$$u_j = x_j \beta + \varepsilon_j \quad j = 1, \ldots, m$$

$$[\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m] \sim N(0, \Sigma)$$

also abhängigen Fehlertermen angebracht. Dies führt zu mehrdimensionalen Integralen, die mit „simulation estimators“ zu lösen sind (vgl. Kapitel III.7).

Ist die IIA-Annahme nicht erfüllt, dann kann neben dem MULTIVARIATEN LOGIT/PROBIT Modell auch ein NESTED LOGIT Modell oder ein HETEROSKEDASTISCHES LOGIT Modell verwendet werden, die im nächsten Abschnitt behandelt werden.

7 LIMDEP: Marginale Effekte und vorhergesagte Wahrscheinlichkeiten

Zur Berechnung der marginalen Effekte und der vorhergesagten Wahrscheinlichkeiten für das MNL-Modell sind in Tabelle 3.5a und Tabelle 3.5b die entsprechenden LIMDEP-Programme aufgeführt. Dabei werden die marginalen Effekte, also die Änderung der Wahrscheinlichkeit für die Wahl einer Alternative bei Änderung einer Einheit von x, prinzipiell an den Mittelwerten der erklärenden Variablen x vorgenommen. Als Mittelwert bei Dummies wird die relative Häufigkeit der Einsen (relativ zu allen Beobachtungen) eingesetzt.

Tab. 3.5a: Marginal Effects (LIMDEP)

```
?=============================================================================
? Marginal effects for the DISCRETE CHOICE model.                           =
? All marginal effects are a multiple of the coefficient vector.            =
? dPj/dxl =  r(j,l) * beta                                                  =
? The following computes the matrix of scale factors r(j,l)                 =
? User must define the data matrix with the following line:                 =
?                                                                           =
?       NAMELIST          ; X = ... $                                       =
?                                                                           =
? User's dependent variable is Y.  Change the following line               =
?                                                                           =
?       CREATE            ; Y = the dependent variable $                    =
?                                                                           =
? User must also modify DISC command below for the choice set.              =
?=============================================================================
? Estimate model first. This is the only command that must be set by the user.
?
       DISC              ;  Lhs = Y ; Rhs = X
                         ;  Choices = ... $
?
? Pick up number of regressors, number of probabilities.
?
       CALC              ;  k = col(X) ; J = max(choice)
                         ;  minusJ = -J ; sumebx = 0 $
?
? Creates indicator of choice, 1,2,3,4,1,2,3,4,...
?
       CREA              ;  choice = trn(minusJ,0) $
?
? Set up for a matrix of means of x's by the J choices.
?
       MATRIX            ;  means=init(j,k,0) ; ones=init(j,1,1) $
?
? The procedure will loop over the J choices,
? injecting vectors into the matrix
?
       PROCEDURE
          REJECT         ; New ; choice # i $
          MATRIX         ; xbr = mean(X) ; means(i,*) = xbr $
       ENDPROC
       EXECUTE           ;  i=1,J $
?
? Compute vector of probabilities at means of regressors
?
       MATRIX  ;  ebx  = means | b
               ;  ebx  = expn(ebx) ; sumebx=ebx' | ones $
       MATRIX  ;  prob = 1/sumebx * ebx $
?
? Double loop to fill cells of matrix using the probabilities.  Then display.
?
       MATRIX            ;  scale=init(j,j,0) $
       PROCEDURE
              CALC       ;  l=1  $
              LABEL      ;  10  $
              CALC       ;  m=1 $
              LABEL      ;  20 $
              CALC       ;  me=eql(l,m)*prob(l)-prob(l)*prob(m) $
              MATRIX     ;  scale(l,m)=me $
              CALC       ;  m=m+1 $
              GO TO      ;  20 ; m <= j $
              CALC       ;  l=l+1 $
              GO TO      ;  10 ; l <= j$
       ENDPROCEDURE
       EXECUTE
       MATRIX            ;  Scale$
```

Tab. 3.5b: Predicted Probalities (LIMDEP)

```
?=====================================================================
? Predicted probabilities for the multinomial logit model           =
?                                                                   =
? User must define the data matrix with the following line:         =
?                                                                   =
?       NAMELIST          ; X = ... $                               =
?                                                                   =
? User's dependent variable is Y.  Change the following line       =
?                                                                   =
?       CREATE            ; Y = the dependent variable $            =
?                                                                   =
? User must set up columns for predicted probabilities with:        =
?       CREATE       ; P0 = 0 ; P1 = 0 ; ... ; PJ = 0 $             =
?       NAMELIST     ; Probs = P0, P1, ..., PJ $                    =
?=====================================================================
?
      LOGIT         ; Lhs = y    ; Rhs = X $
?
? Problem size and some counters
?
      CALC          ; Nolist              ; K = col(X)
                    ; jplus1 = col(Probs) ; i = 2 ; i1=1  ; i2=K $
?
? Parameters are in a vector á. Arrange in a matrix [á1,á2,...]
?
      MATRIX        ; Nolist ; ones = init(jplus1,1,1)
                             ; beta=init (k,jplus1,0) $
      LABEL         ; 100 $
      MATRIX        ; Nolist  ; bj=part(b,i1,i2) ;beta(*,i)=bj $
      CALC          ; Nolist  ; i1=i1+k;i2=i2+k;i=i+1 $
      GO TO         ; 100 ; i <= jplus1 $
?
? Now assemble fitted probabilities. Note use of a loop over variables
?
      MATRIX        ; Probs = Indx(X,Beta) $
      PROCEDURE
            CREATE        ; :Pj = exp(:Pj) $
      ENDPROC
      EXECUTE             ; :Pj = Probs $
      CREATE              ; Sum = DOT(Probs,ones) $
      PROCEDURE
             CREATE       ; :Pj = :Pj / Sum $
      ENDPROC
      EXECUTE             ; :Pj = Probs $
```

8 Beispiele multipler Entscheidungsmodelle

8.1 MULTINOMIALES LOGIT / LIMDEP: Berufliche Stellung

Occupation als Funktion sozioökonomischer Variablen (in Anlehnung an Schmidt und Strauss 1975a, eigene fiktive Daten)

$$P_j = \frac{\exp(x_i \beta_j)}{1 + \sum_{k=1}^{J} \exp(x_i \beta_k)} \qquad j = 1, \ldots, J = m$$

Alternativen:

Occupation

y = 0	Menial (niedrige Dienste)	y = j
1	blue collar	
2	white collar	
3	professional (Freie Berufe)	

Regressoren:

x = (1, education, sex, income)
Normierung auf j = 0 (menial)

Ergebnis:

$$\beta_j = \beta_{0j} + \beta_{\text{education},j} + \ldots + \beta_{\text{income},j}$$

also alternativenspezifisch (d.h. es gibt für jede Alternative einen Ergebnisvektor)

LIMDEP-Ergebnisse: MNL - Berufliche Stellung

```
/* MULTINOMIAL LOGIT Beispiel
   - mit marginalen Effekten
   - und Hausman und McFadden (1984) IIA-Test
-------------------------------------------------------
Schmidt und Strauss 1975a,b, eigene fiktive Daten

Alternativen: Occupation
      y=0 menial
        1 blue collar
        2 white collar
        3 professional

Regressors:
      x=(educat (in years),sex (0=female,1=male),income)
-------------------------------------------------------
*/
--> read;nobs=30;nvar=4;names=1;file=mnlexam-dat.txt$
--> list;y,educat,sex,income;file$
```

Listing of raw data (Current sample)

Line	Observ.	Y	EDUCAT	SEX	INCOME
1	1	0	7	1	1100
2	2	0	9	1	1700
3	3	1	8	1	1400
4	4	1	7	1	1300
5	5	2	12	0	2850
6	6	3	13	1	3800
7	7	2	10	1	2700
8	8	2	12	0	2810
9	9	3	10	0	2950
10	10	3	14	0	4500
11	11	2	13	1	3000
12	12	0	9	0	800
13	13	0	8	0	750
14	14	2	9	0	2400
15	15	3	11	1	3400
16	16	2	10	1	2600
17	17	1	9	0	1900
18	18	1	8	0	1850
19	19	3	12	1	3400
20	20	3	11	0	3300
21	21	0	9	0	1200
22	22	0	8	1	1900
23	23	2	10	1	1900
24	24	2	10	0	1600
25	25	3	9	1	3200
26	26	3	7	0	2200
27	27	0	8	0	1425
28	28	0	9	0	1655
29	29	1	11	1	2000
30	30	3	12	0	3100

```
--> dstat; rhs=*$
```

Descriptive Statistics
All results based on nonmissing observations.

Variable	Mean	Std.Dev.	Minimum	Maximum	Cases
All observations in current sample					
Y	1.60000000	1.19192687	.000000000	3.00000000	30
EDUCAT	9.83333333	1.91335296	7.00000000	14.0000000	30
SEX	.466666667	.507416263	.000000000	1.00000000	30
INCOME	2289.66667	939.892633	750.000000	4500.00000	30

```
--> MLOGIT; lhs=y
    ; rhs=educat,sex,income
    ; marginal effects$
```

Normal exit from iterations. Exit status=0.

Multinomial Logit Model	
Maximum Likelihood Estimates	
Model estimated: Dec 23, 2005 at 10:11:39AM.	
Dependent variable	Y
Weighting variable	None
Number of observations	30
Iterations completed	10
Log likelihood function	-19.03155
Restricted log likelihood	-40.94265
Chi squared	43.82218
Degrees of freedom	6
Prob[ChiSqd > value] =	.0000000

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]	Mean of X
	Characteristics in numerator of Prob[Y = 1]				
EDUCAT	-.5275697122	.36983600	-1.426	.1537	9.8333333
SEX	.3272823579	1.2711104	.257	.7968	.46666667
INCOME	.2555341358E-02	.20085310E-02	1.272	.2033	2289.6667
	Characteristics in numerator of Prob[Y = 2]				
EDUCAT	-1.005280554	.45493951	-2.210	.0271	9.8333333
SEX	-.5304989144	1.3321542	-.398	.6905	.46666667
INCOME	.5294018199E-02	.22833436E-02	2.319	.0204	2289.6667
	Characteristics in numerator of Prob[Y = 3]				
EDUCAT	-5.286995577	2.7788625	-1.903	.0571	9.8333333
SEX	-3.800175713	3.2381911	-1.174	.2406	.46666667
INCOME	.2183391192E-01	.10775934E-01	2.026	.0427	2289.6667

(Note: E+nn or E-nn means multiply by 10 to + or -nn power.)

Information Statistics for Discrete Choice Model.	M=Model	MC=Constants Only	M0=No Model
Criterion F (log L)	-19.03155	-40.94265	-41.58883
LR Statistic vs. MC	43.82218	.00000	.00000
Degrees of Freedom	6.00000	.00000	.00000
Prob. Value for LR	.00000	.00000	.00000
Entropy for probs.	19.03155	40.94265	41.58883
Normalized Entropy	.45761	.98446	1.00000
Entropy Ratio Stat.	45.11455	1.29237	.00000
Bayes Info Criterion	58.47029	102.29248	103.58485
BIC - BIC(no model)	45.11455	1.29237	.00000
Pseudo R-squared	.53517	.00000	.00000
Pct. Correct Prec.	66.66667	.00000	25.00000

Means:	y=0	y=1	y=2	y=3	yu=4	y=5,	y=6	y>=7
Outcome	.2667	.1667	.2667	.3000	.0000	.0000	.0000	.0000
Pred.Pr	.2533	.1572	.2821	.3075	.0000	.0000	.0000	.0000

Notes: Entropy computed as Sum(i)Sum(j)Pfit(i,j)*logPfit(i,j).
Normalized entropy is computed against M0.
Entropy ratio statistic is computed against M0.
BIC = 2*criterion - log(N)*degrees of freedom.
If the model has only constants or if it has no constants, the statistics reported here are not useable.

Partial derivatives of probabilities with respect to the vector of characteristics.
They are computed at the means of the Xs.
Observations used for means are All Obs.
A full set is given for the entire set of outcomes, Y = 0 to Y = 3.
Probabilities at the mean vector are
0= .094 1= .214 2= .690 3= .002

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]	Mean of X
	Marginal effects on Prob[Y = 0]				
EDUCAT	.7721091130E-01	.55472538E-01	1.392	.1640	9.8333333
SEX	.2874179797E-01	.10102507	.285	.7760	.46666667
INCOME	-.4008788098E-03	.25208351E-03	-1.590	.1118	2289.6667
	Marginal effects on Prob[Y = 1]				
EDUCAT	.6194091056E-01	.67029407E-01	.924	.3554	9.8333333
SEX	.1348642002	.19564098	.689	.4906	.46666667
INCOME	-.3608391504E-03	.28786010E-03	-1.254	.2100	2289.6667
	Marginal effects on Prob[Y = 2]				
EDUCAT	-.1294216575	.84826868E-01	-1.526	.1271	9.8333333
SEX	-.1559952886	.24189957	-.645	.5190	.46666667
INCOME	.7234259335E-03	.35974406E-03	2.011	.0443	2289.6667
	Marginal effects on Prob[Y = 3]				
EDUCAT	-.9730164410E-02	.31746457E-01	-.306	.7592	9.8333333
SEX	-.7610709536E-02	.25445217E-01	-.299	.7649	.46666667
INCOME	.3829202667E-04	.12566458E-03	.305	.7606	2289.6667

```
 (Note: E+nn or E-nn means multiply by 10 to + or -nn power.)


Marginal Effects Averaged Over Individuals
--------+------+------+------+------+
Variable|  Y=00|  Y=01|  Y=02|  Y=03|
--------+------+------+------+------+
EDUCAT  | .0780| .0068| .0347|-.1195|
SEX     | .0066| .0752| .0107|-.0924|
INCOME  |-.0004| .0000| .0000| .0005|
--------+------+------+------+------+


Averages of Individual Elasticities of Probabilities
--------+------+------+------+------+
Variable|  Y=00|  Y=01|  Y=02|  Y=03|
--------+------+------+------+------+
EDUCAT  |******|******|******|******|
SEX     | .5447| .6974| .2971|******|
INCOME  |******|******|******|******|
--------+------+------+------+------+

Frequencies of actual & predicted outcomes
Predicted outcome has maximum probability.

            Predicted
------  --------------------  +  -----
Actual       0    1    2    3  |  Total
------  --------------------  +  -----
  0          7    0    1    0  |      8
  1          3    0    2    0  |      5
  2          2    0    5    1  |      8
  3          0    0    1    8  |      9
------  --------------------  +  -----
Total       12    0    9    9  |     30
```

? Abspeichern des unrestringierten Modells (mit professionals)

```
--> Matrix;list;bun=b(1:6);vu=varb(1:6,1:6)$

Matrix BUN      has  6 rows and  1 columns.
                 1
         +--------------
       1|     -.5276
       2|      .3273
       3|      .0026
       4|    -1.0053
       5|     -.5305
       6|      .0053
```

Matrix: VU
[6,6]

? Berechnen des restringierten Modells (ohne professionals)

```
--> reject;y=3$
--> MLOGIT; lhs=y
    ; rhs=educat,sex,income
    ; marginal effects$

Normal exit from iterations. Exit status=0.

+---------------------------------------------+
| Multinomial Logit Model                     |
| Maximum Likelihood Estimates                |
| Model estimated: Dec 23, 2005 at 10:11:40AM.|
```

```
| Dependent variable                Y     |
| Weighting variable             None     |
| Number of observations           21     |
| Iterations completed              6     |
| Log likelihood function      -16.41780  |
| Restricted log likelihood    -22.61672  |
| Chi squared                   12.39783  |
| Degrees of freedom                4     |
| Prob[ChiSqd > value] =     .1462554E-01  |
+-----------------------------------------+
+---------+-------------+----------------+--------+---------+----------+
|Variable | Coefficient | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+-------------+----------------+--------+---------+----------+
         Characteristics in numerator of Prob[Y = 1]
 EDUCAT     -.5261516852       .36923037   -1.425   .1542     9.3333333
 SEX         .3290315903      1.2716506     .259   .7958     .47619048
 INCOME    .2548521873E-02  .20067390E-02   1.270   .2041     1849.5238
         Characteristics in numerator of Prob[Y = 2]
 EDUCAT     -1.021397444       .45630348   -2.238   .0252     9.3333333
 SEX        -.5437659508      1.3353033     -.407   .6838     .47619048
 INCOME    .5383003436E-02  .22858476E-02   2.355   .0185     1849.5238
 (Note: E+nn or E-nn means multiply by 10 to + or -nn power.)
```

Matrix: Las
[6,4]

```
+-------------------------------------------------------------------------+
| Information Statistics for Discrete Choice Model.                       |
|                           M=Model MC=Constants Only   M0=No Model       |
| Criterion F (log L)      -16.41780         -22.61672     -23.07086      |
| LR Statistic vs. MC       12.39783            .00000        .00000      |
| Degrees of Freedom         4.00000            .00000        .00000      |
| Prob. Value for LR          .01463            .00000        .00000      |
| Entropy for probs.        16.41780          22.61672      23.07086      |
| Normalized Entropy          .71163            .98032       1.00000      |
| Entropy Ratio Stat.       13.30612            .90828        .00000      |
| Bayes Info Criterion      45.01369          57.41152      58.31981      |
| BIC - BIC(no model)       13.30612            .90828        .00000      |
| Pseudo R-squared            .27409            .00000        .00000      |
| Pct. Correct Prec.        61.90476            .00000      33.33333      |
| Means:       y=0    y=1    y=2    y=3    yu=4   y=5,    y=6   y>=7      |
| Outcome     .3810  .2381  .3810  .0000  .0000  .0000  .0000  .0000      |
| Pred.Pr     .3620  .2254  .4126  .0000  .0000  .0000  .0000  .0000      |
| Notes: Entropy computed as Sum(i)Sum(j)Pfit(i,j)*logPfit(i,j).          |
|        Normalized entropy is computed against M0.                       |
|        Entropy ratio statistic is computed against M0.                  |
|        BIC = 2*criterion - log(N)*degrees of freedom.                   |
|        If the model has only constants or if it has no constants,       |
|        the statistics reported here are not useable.                    |
+-------------------------------------------------------------------------+
```

```
+-----------------------------------------+
| Partial derivatives of probabilities with |
| respect to the vector of characteristics. |
| They are computed at the means of the Xs. |
| Observations used for means are All Obs.  |
| A full set is given for the entire set of |
| outcomes, Y        = 0 to Y        =  2.  |
| Probabilities at the mean vector are      |
|  0= .319 1= .306 2= .375                  |
+-----------------------------------------+
+---------+-------------+----------------+--------+---------+----------+
|Variable | Coefficient | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+-------------+----------------+--------+---------+----------+
         Marginal effects on Prob[Y = 0]
 EDUCAT       .1734663389   .80558108E-01    2.153   .0313     9.3333333
 SEX       .3296481656E-01      .24622957     .134   .8935     .47619048
```

```
 INCOME   -.8923297678E-03  .38353461E-03   -2.327   .0200     1849.5238
         Marginal effects on Prob[Y = 1]
 EDUCAT    .5592146622E-02  .66207126E-01     .084   .9327     9.3333333
 SEX           .1323296968       .22535631     .587   .5571     .47619048
 INCOME   -.7712815255E-04  .32103757E-03    -.240   .8101     1849.5238
         Marginal effects on Prob[Y = 2]
 EDUCAT      -.1790584855   .79835953E-01   -2.243   .0249     9.3333333
 SEX         -.1652945133       .27094150    -.610   .5418     .47619048
 INCOME    .9694579204E-03  .39190350E-03    2.474   .0134     1849.5238
 (Note: E+nn or E-nn means multiply by 10 to + or -nn power.)


Marginal Effects Averaged Over Individuals
--------+------+------+------+
Variable|  Y=00|  Y=01|  Y=02|
--------+------+------+------+
EDUCAT  | .1092|-.0039|-.1053|
SEX     | .0075| .0975|-.1050|
INCOME  |-.0006| .0000| .0006|
--------+------+------+------+


Averages of Individual Elasticities of Probabilities
--------+------+------+------+
Variable|  Y=00|  Y=01|  Y=02|
--------+------+------+------+
EDUCAT  |5.2602| .3495|******|
SEX     | .0566| .2133|-.2024|
INCOME  |******|******|3.8397|
--------+------+------+------+

Frequencies of actual & predicted outcomes
Predicted outcome has maximum probability.
            Predicted
------  ---------------  +  -----
Actual       0    1    2  |  Total
------  ---------------  +  -----
  0          7    0    1  |      8
  1          3    0    2  |      5
  2          2    0    6  |      8
------  ---------------  +  -----
Total       12    0    9  |     21



?-----------------------------------
? Hausman and McFadden Test auf IIA
?-----------------------------------


--> Matrix;list
    ;br=b(1:6)
    ;bu=bun(1:6)
    ;db=br-bu
    ;vr=varb(1:6,1:6)
    ;H=db'* nvsm(vr,-vu) * db$


Matrix BR       has  6 rows and  1 columns.
                 1
        +--------------
       1|     -.5262
       2|      .3290
       3|      .0025
       4|    -1.0214
       5|     -.5438
       6|      .0054

Matrix BU       has  6 rows and  1 columns.
                 1
        +--------------
```

```
       1|     -.5276
       2|      .3273
       3|      .0026
       4|    -1.0053
       5|     -.5305
       6|      .0053

Matrix DB       has  6 rows and  1 columns.
                 1
        +--------------
       1|      .0014
       2|      .0017
       3| -.6819485D-05
       4|     -.0161
       5|     -.0133
       6|  .8898524D-04
```

Matrix: VR
[6,6]

Da die Differenz zwischen den geschätzen Koeffizienten des unrestringierten Modells (mit professionals geschätzt) und den Koeffizienten des restringierten Modells (ohne professionals geschätzt) doch sehr gering ist, wird wohl die zusätzliche Alternative (professionals) keinen signifikanten Einfluß auf die Wahl der anderen Alternativen haben; das wird nun mit dem Hausmann-Test getestet:

```
Matrix H        has  1 rows and  1 columns.
                 1
        +--------------
       1|      .4659
```

mit Anzahl der Freiheitsgrade (DF) gleich der Anzahl der geschätzten Koeffizienten (Zeilen) = 6

```
--> calc; list
    ; tabchi95=ctb(.95,6)
    ; tabchi99=ctb(.99,6)
    ; chi(H,(Row(db)))
    ; pvalue = 1-chi(H,6)$
    TABCHI95=  .12591587245499980D+02
    TABCHI99=  .16811893831400030D+02
    Result  =  .17709056210358640D-02
    PVALUE  =  .99822909437896410D+00
Calculator: Computed   4 scalar results
```

Resultat:

mit einem Chi2-Wert von H=.4659 < 12,6 (tabchi95) bzw. 16,8 (tabchi99) bzw.

mit einem p-value von .9982 > α=0,05 bzw. α=0,01

ist die IIA-Nullhypothese, eine Alternative hat keinen Einfluss, nicht abzulehnen.

D.h. die IIA-Annahme (Independence of Irrelevant Alternatives) gilt, das Modell passt, d.h. eine zusätzliche Alternative (hier professionals) beeinflußt das Wahrscheinlichkeitsverhältnis zwischen den anderen Alternativen nicht.

8.2 MULTINOMIALES LOGIT / Stata: Berufliche Stellung

```
. * mlogit / stata
. * hausman IIA test
. *
. * j.merz 27.5.2013
.
. clear all
```

```
. set mem 500m

Current memory allocation

                    current                                  memory usage
    settable          value     description                  (1M = 1024k)
    ---------------------------------------------------------------------
    set maxvar         5000     max. variables allowed           1.909M
    set memory          500M    max. data space                500.000M
    set matsize         400     max. RHS vars in models          1.254M
                                                             -----------
                                                                503.163M

. capture log close

. set more off

.
. log using "mlogit_occupation_hausman_test_skript.log", replace
-------------------------------------------------------------------------------------
---------------------------------------
      name:  <unnamed>
       log:
X:\EWF\_Discrete Choice\Beispiele\MNL estimates
Stata\mlogit_occupation_hausman_test_skript.log
  log type:  text
 opened on:  30 May 2013, 09:56:16

. use MNLEXAMPLE-dat_V10, replace

.
. sum *

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
      observ |        30        15.5    8.803408          1         30
           y |        30         1.6    1.191927          0          3
      educat |        30    9.833333    1.913353          7         14
         sex |        30    .4666667    .5074163          0          1
      income |        30    2289.667    939.8926        750       4500

. global x educat sex

.
.
. * MNL
. mlogit y $x, base(0)

Iteration 0:   log likelihood = -40.942646
Iteration 1:   log likelihood =  -33.14412
Iteration 2:   log likelihood =  -32.64134
Iteration 3:   log likelihood = -32.628838
Iteration 4:   log likelihood = -32.628809
Iteration 5:   log likelihood = -32.628809

Multinomial logistic regression                   Number of obs   =         30
                                                  LR chi2(6)      =      16.63
                                                  Prob > chi2     =     0.0108
Log likelihood = -32.628809                       Pseudo R2       =     0.2031

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
0            |  (base outcome)
-------------+----------------------------------------------------------------
1            |
      educat |   .2136465   .5149648     0.41   0.678     -.795666    1.222959
         sex |   .9595302   1.180556     0.81   0.416    -1.354318    3.273378
       _cons |  -2.748282   4.510059    -0.61   0.542    -11.58783    6.091271
-------------+----------------------------------------------------------------
2            |
      educat |   1.248257   .5363583     2.33   0.020     .1970145      2.2995
         sex |   .3959491   1.236407     0.32   0.749    -2.027365    2.819263
       _cons |  -11.86602   4.967337    -2.39   0.017    -21.60182   -2.130219
```

```
-------------+----------------------------------------------------------------
3            |
      educat |   1.343701   .5423721     2.48   0.013     .2806712    2.406731
         sex |   .1559369   1.247926     0.12   0.901    -2.289953    2.601827
       _cons |  -12.67277   5.038821    -2.52   0.012    -22.54868   -2.796859
------------------------------------------------------------------------------

. estimates store mall

.
. mlogit y $x if y!=1, base(0)

Iteration 0:   log likelihood =  -27.42581
Iteration 1:   log likelihood =  -21.51493
Iteration 2:   log likelihood = -20.667397
Iteration 3:   log likelihood = -20.628503
Iteration 4:   log likelihood = -20.628373
Iteration 5:   log likelihood = -20.628373

Multinomial logistic regression                 Number of obs   =         25
                                                LR chi2(4)      =      13.59
                                                Prob > chi2     =     0.0087
Log likelihood = -20.628373                     Pseudo R2       =     0.2478

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
0            |  (base outcome)
-------------+----------------------------------------------------------------
2            |
      educat |    1.41404   .6590666     2.15   0.032     .1222933    2.705787
         sex |   .4192728   1.291989     0.32   0.746    -2.112979    2.951524
       _cons |  -13.42229   6.082738    -2.21   0.027    -25.34424    -1.50034
-------------+----------------------------------------------------------------
3            |
      educat |   1.508813   .6652663     2.27   0.023     .2049151    2.812711
         sex |   .1944777    1.30358     0.15   0.881    -2.360493    2.749448
       _cons |  -14.22893   6.155592    -2.31   0.021    -26.29367   -2.164192
------------------------------------------------------------------------------

. estimates store m1

.
. mlogit y $x if y!=2, base(0)

Iteration 0:   log likelihood = -23.545191
Iteration 1:   log likelihood = -17.920728
Iteration 2:   log likelihood = -17.823409
Iteration 3:   log likelihood = -17.822758
Iteration 4:   log likelihood = -17.822758

Multinomial logistic regression                 Number of obs   =         22
                                                LR chi2(4)      =      11.44
                                                Prob > chi2     =     0.0220
Log likelihood = -17.822758                     Pseudo R2       =     0.2430

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
0            |  (base outcome)
-------------+----------------------------------------------------------------
1            |
      educat |   .1828981   .4759065     0.38   0.701    -.7498616    1.115658
         sex |   .9545121   1.179428     0.81   0.418    -1.357125    3.266149
       _cons |  -2.484664   4.178924    -0.59   0.552     -10.6752    5.705877
-------------+----------------------------------------------------------------
3            |
      educat |   1.083485   .4874492     2.22   0.026     .1281023    2.038868
         sex |   .0439511   1.317416     0.03   0.973    -2.538136    2.626038
       _cons |  -10.15318   4.504288    -2.25   0.024    -18.98142   -1.324934
------------------------------------------------------------------------------

. estimates store m2
```

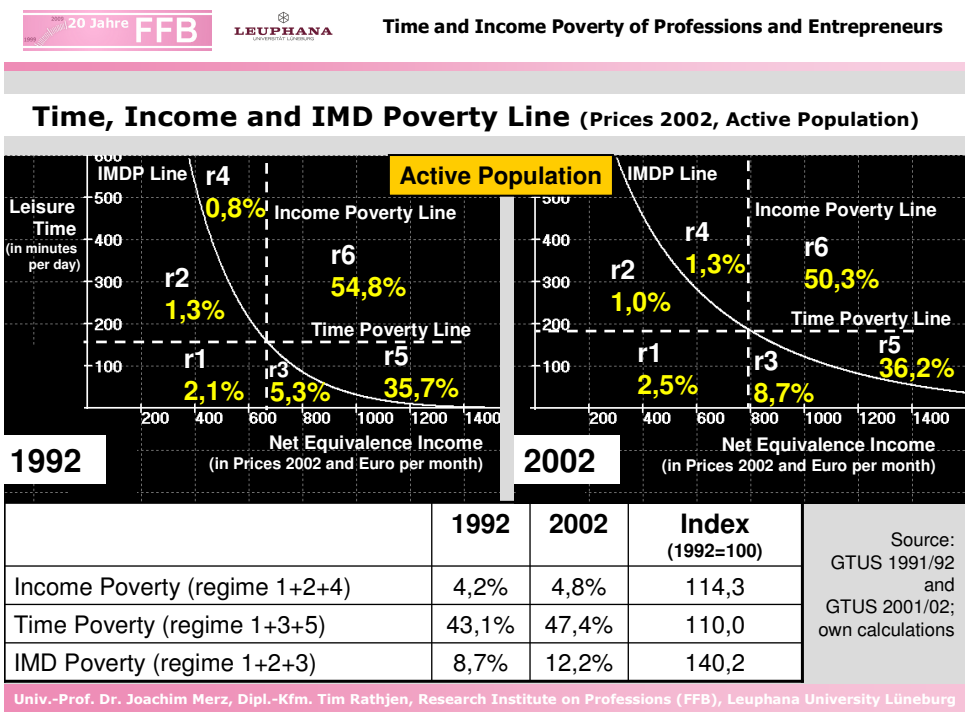
```
.
. mlogit y $x if y!=3, base(0)

Iteration 0:   log likelihood = -22.616717
Iteration 1:   log likelihood = -15.803874
Iteration 2:   log likelihood = -15.448958
Iteration 3:   log likelihood = -15.443937
Iteration 4:   log likelihood = -15.443928
Iteration 5:   log likelihood = -15.443928

Multinomial logistic regression                 Number of obs   =         21
                                                LR chi2(4)      =      14.35
                                                Prob > chi2     =     0.0063
Log likelihood = -15.443928                     Pseudo R2       =     0.3171

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
0            |  (base outcome)
-------------+----------------------------------------------------------------
1            |
      educat |   .3901843   .6487628     0.60   0.548    -.8813674    1.661736
         sex |   1.085755   1.223891     0.89   0.375    -1.313027    3.484536
       _cons |  -4.300136   5.703116    -0.75   0.451    -15.47804    6.877766
-------------+----------------------------------------------------------------
2            |
      educat |   2.062776   1.018718     2.02   0.043      .066126    4.059426
         sex |   .2306714   1.599457     0.14   0.885    -2.904207    3.365549
       _cons |  -19.33526   9.288365    -2.08   0.037    -37.54012   -1.130398
------------------------------------------------------------------------------

. estimates store m3

.
. hausman m1 mall, alleqs constant

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |       m1          mall        Difference          S.E.
-------------+------------------------------------------------------------
2            |
      educat |     1.41404     1.248257        .1657826        .3829994
         sex |    .4192728     .3959491        .0233237        .3748757
       _cons |   -13.42229    -11.86602       -1.556267        3.510735
-------------+------------------------------------------------------------
3            |
      educat |    1.508813     1.343701        .1651121        .3852424
         sex |    .1944777     .1559369        .0385409         .376831
       _cons |   -14.22893    -12.67277       -1.556163        3.535759
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from mlogit
            B = inconsistent under Ha, efficient under Ho; obtained from mlogit

    Test:  Ho:  difference in coefficients not systematic

                  chi2(6) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =        0.23
                Prob>chi2 =      0.9998
                (V_b-V_B is not positive definite)

. hausman m2 mall, alleqs constant

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |       m2          mall        Difference          S.E.
-------------+------------------------------------------------------------
1            |
      educat |    .1828981     .2136465       -.0307484               .
         sex |    .9545121     .9595302       -.0050181               .
       _cons |   -2.484664    -2.748282        .2636181               .
-------------+------------------------------------------------------------
3            |
```

```
      educat |    1.083485     1.343701       -.2602157              .
         sex |    .0439511     .1559369       -.1119857        .4222137
       _cons |   -10.15318    -12.67277        2.519592              .
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from mlogit
            B = inconsistent under Ha, efficient under Ho; obtained from mlogit

    Test:  Ho:  difference in coefficients not systematic

                  chi2(6) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =    -1.51    chi2<0 ==> model fitted on these
                                        data fails to meet the asymptotic
                                        assumptions of the Hausman test;
                                        see suest for a generalized test

. hausman m3 mall, alleqs constant

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |       m3          mall        Difference          S.E.
-------------+----------------------------------------------------------------
1            |
      educat |    .3901843     .2136465        .1765379         .394594
         sex |    1.085755     .9595302        .1262245        .3227919
       _cons |   -4.300136    -2.748282       -1.551855        3.490689
-------------+----------------------------------------------------------------
2            |
      educat |    2.062776     1.248257        .8145187        .8660864
         sex |    .2306714     .3959491       -.1652778        1.014672
       _cons |   -19.33526    -11.86602       -7.469238        7.848521
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from mlogit
            B = inconsistent under Ha, efficient under Ho; obtained from mlogit

    Test:  Ho:  difference in coefficients not systematic

                  chi2(6) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =        1.33
                Prob>chi2 =      0.9700
                (V_b-V_B is not positive definite)

.
end of do-file
```

8.3 MULTINOMIALES LOGIT: Sind Selbständige zeit- und einkommensarm?

Merz, Joachim und Tim Rathjen, 2010, Sind Selbständige zeit- und einkommensarm? Eine Mikroanalyse der Dynamik interdependenter multidimensionaler Armut mit dem Sozio-ökonomischen Panel und den deutschen Zeitbudgeterhebungen, FFB-Diskussionspapier Nr. 82, Lüneburg

… Selbständige sind überproportional von gleichzeitiger Zeit- und Einkommensarmut betroffen. Darüber hinaus ist ein erheblicher Anteil der zeit- aber nicht einkommensarmen Selbständigen nicht in der Lage, ihr Zeitdefizit durch Einkommen zu kompensieren.



	1992	2002	Index (1992=100)
Income Poverty (regime 1+2+4)	4,2%	4,8%	114,3
Time Poverty (regime 1+3+5)	43,1%	47,4%	110,0
IMD Poverty (regime 1+2+3)	8,7%	12,2%	140,2

Source: GTUS 1991/92 and GTUS 2001/02; own calculations

Univ.-Prof. Dr. Joachim Merz, Dipl.-Kfm. Tim Rathjen, Research Institute on Professions (FFB), Leuphana University Lüneburg

Tabelle 6: Ergebnisse von Zeit-, Einkommens- und interdependenter multidimensionaler Armut - Erwerbstätige

Multinomiales Logit Referenz: Sektor 6 (keine Zeit- oder Einkommensarmut) Std. Err. adjusted for 4146 Clusters	**Sektor 1**		**Sektor 2**		**Sektor 3**		**Sektor 4**		**Sektor 5**	
	1992	**2002**	**1992**	**2002**	**1992**	**2002**	**1992**	**2002**	**1992**	**2002**
Konstante	6,513	8,801	14,137	16,246	-7,097	-5,963	11,092	14,180	-5,746	-5,676
	(0,013)	(0,017)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,001)	(0,000)	(0,000)
persönliche Charakteristika										
weiblich	1,035	1,322	0,361	-0,560	1,075	1,028	-0,611	-0,094	0,843	0,699
	(0,002)	(0,008)	(0,312)	(0,336)	(0,000)	(0,000)	(0,103)	(0,841)	(0,000)	(0,000)
Alter	0,180	0,007	-0,115	-0,232	0,064	0,084	0,055	-0,903	0,0818	0,084
	(0,061)	(0,946)	(0,908)	(0,051)	(0,104)	(0,012)	(0,600)	(0,378)	(0,000)	(0,000)
Alter²/100	-0,163	0,046	0,596	0,334	-0,063	-0,085	-0,033	0,120	-0,086	-0,085
	(0,198)	(0,715)	(0,644)	(0,024)	(0,196)	(0,043)	(0,812)	(0,356)	(0,000)	(0,000)
Bildung (Ref.: kein Abschluss)										
Hauptschule	1,722	0,463	1,224	1,623	0,783	-0,411	1,437	0,786	0,004	-0,313
	(0,088)	(0,747)	(0,129)	(0,253)	(0,219)	(0,451)	(0,182)	(0,502)	(0,989)	(0,390)
Realschule	1,172	0,745	1,040	2,224	0,651	-0,494	1,139	0,866	-0,047	-0,281
	(0,207)	(0,596)	(0,159)	(0,115)	(0,306)	(0,364)	(0,272)	(0,463)	(0,880)	(0,442)
Abitur (ohne Hochschulabschluss)	0,769	0,745	0,942	2,225	0,869	-0,432	-31,653	-0,053	0,004	-0,222
	(0,439)	(0,610)	(0,267)	(0,129)	(0,180)	(0,430)	(0,000)	(0,956)	(0,989)	(0,544)
Hochschulabschluss	1,579	2,126	1,757	3,310	0,761	-0,183	1,532	1,093	0,156	0,006
	(0,138)	(0,151)	(0,037)	(0,028)	(0,251)	(0,745)	(0,185)	(0,396)	(0,628)	(0,988)
Berufsstatus (Referenz: Arbeiter)										
Selbständige	0,385		0,538		0,964		0,243		0,450	
	(0,431)		(0,286)		(0,000)		(0,664)		(0,000)	
Freiberufler		1,038		-1,028		0,279		0,808		0,007
		(0,274)		(0,404)		(0,365)		(0,425)		(0,966)
Unternehmer		1,842		1,048		0,784		0,878		0,406
		(0,001)		(0,131)		(0,000)		(0,190)		(0,003)
Beamte	0,581	0,041	-30,011	1,595	0,285	0,001	2,057	-0,227	0,15	0,045
	(0,717)	(0,964)	(0,000)	(0,048)	(0,339)	(0,998)	(0,021)	(0,841)	(0,207)	(0,705)
Angestellte	-0,635	0,125	-0,615	0,500	0,257	-0,286	-0,279	0,205	0,133	0,561
	(0,101)	(0,789)	(0,141)	(0,91)	(0,081)	(0,838)	(0,545)	(0,656)	(0,069)	(0,523)
anderer Berufsstatus	0,757	0,550	0,561	0,438	0,199	-0,050	0,717	0,961	-0,462	-0,260

	(0,227)	(0,463)	(0,376)	(0,601)	(0,536)	(0,825)	(0,231)	(0,222)	(0,709)	(0,071)
Arbeitssituation										
Erwerbsarbeitszeit	0,008	0,010	0,002	0,004	0,009	0,009	-0,000	0,001	0,005	0,005
	(0,000)	(0,000)	(0,088)	(0,025)	(0,000)	(0,000)	(0,767)	(0,554)	(0,000)	(0,000)
Einkommenssituation										
persönliches Nettoeinkommen	-0,109	-0,016	0,011	-0,016	-0,001	-0,002	-0,009	-0,122	-0,000	0,000
	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,436)	(0,125)
Residualeinkommen	-0,017	-0,015	-0,018	-0,017	-0,002	0,001	-0,013	-0,130	0,000	0,000
	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,003)	(0,000)
Nationalität (Referenz: Deutsch)										
Nicht-Deutsch		1,504		1,938		-0,336		0,469		-0,274
		(0,267)		(0,189)		(0,443)		(0,806)		(0,327)
Haushaltsstruktur										
(Referenz: Paar ohne Kinder)										
Single-Haushalt	-6,720	-5,400	-6,741	-4,745	-0,880	-0,892	39,811	-4,469	-0,168	0,113
	(0,000)	(0,000)	(0,000)	(0,002)	(0,002)	(0,001)	(0,000)	(0,000)	(0,154)	(0,390)
Paar mit einem Kind	1,730	6,377	1,930	6,755	1,056	0,888	0,952	4,417	0,274	0,078
	(0,011)	(0,000)	(0,003)	(0,000)	(0,000)	(0,000)	(0,188)	(0,000)	(0,011)	(0,435)
Paar mit zwei Kindern	6,429	9,007	6,020	10,280	1,946	1,689	4,290	6,578	0,335	0,148
	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,001)	(0,135)
Paar mit > zwei Kinder	9,603	15,108	9,375	16,425	3,070	2,891	5,678	11,512	0,606	0,137
	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,327)
Alleinerziehende, ein Kind	-0,379	0,520	-1,473	1,122	-0,211	0,272	-0,399	0,225	0,600	0,014
	(0,585)	(0,437)	(0,064)	(0,309)	(0655)	(0,301)	(0,964)	(0,763)	(0,012)	(0,929)
Alleinerziehende, > ein Kind	3,593	4,976	2,787	5,868	1,57	1,172	1,852	3,764	0,309	0,192
	(0,000)	(0,000)	(0,002)	(0,000)	(0,000)	(0,000)	(0,042)	(0,000)	(0,158)	(0,316)
anderer Haushalt	3,28	8,477	3,517	8,650	1,393	1,990	2,478	5,139	0,083	0,121
	(0,000)	(0,001)	(0,000)	(0,001)	(0,000)	(0,000)	(0,000)	(0,028)	(0,359)	(0,506)
Region (Ref.: West-Deutschland)										
Ost-Deutschland	0,586	0,251	0,147	-0,047	0,175	0,319	-0,053	0,003	0,234	0,340
	(0,168)	(0,487)	(0,723)	(0,913)	(0,207)	(0,012)	(0,917)	(0,993)	(0,001)	(0,000)

	1992	**2002**
N (Beobachtungen)	9918	7966
Prob > chi2	0,000	0,000
Pseudo R2 (McFadden)	0,265	0,256

p-value in Klammern .

Quelle: ZBE 1991/92 und 2001/02; eigene Berechnungen; Erwerbstätige.

8.4 CONDITIONAL LOGIT: Travel mode choice

Travel mode choice Sydney/Melbourne (Hensher 1986 nach Greene 1993, S. 668 ff)

$$P_j = \frac{\exp(x_{ij}\beta)}{\sum_{k=1}^{m}\exp(x_{ik}\beta)}$$ keine Normierung, ohne Konstante!

Alternativen: m=4

Regressoren: 10

Ergebnis: $\beta = (\beta_1,\ldots,\beta_{10})'$ also identisch für alle Alternativen

Siehe Tab. 3.6b CONDITIONAL LOGIT: Travel mode choice. Weitere Beispiele (mit alternativen- und individuenspezifischen Charakteristika): Maddala 1983, S. 42 ff, Andreß, Hagenars und Kühnel 1997.

Tab. 3.6b: CONDITIONAL LOGIT: Travel mode choice

Hensher (1986) estimated a model of mode choice for urban travel for a sample of Sydney commuters. The four choices were Car/Driver (C/D), Car/Passenger, Train, and Bus. For a basic model, the attributes were (1) a Car/Driver dummy, (2) a Car/Passenger dummy, (3) a Train dummy, (4) In-Vehicle Time (in minutes), (5) Waiting Time (in minutes), (6) Walking Time (in minutes), (7) In-Vehicle Costs, (8) Parking Cost, (9) Number of Household Business Vehicles Required, and (10) Percentage of Travel Cost Covered by a Nonhousehold Source [(9) and (10) for C/D only]. The sample consists of 1455 observations. A summary of the data is given in the following Table:

	In-Vehicle Cost	In-Vehicle Time	Walk Time	Wait Time	Number
C/D	64,56	28.65	0.76	0.15	953
C/P	4.37	28.32	0.71	2.89	78
Train	98.23	43.84	10.50	8.37	279
Bus	81.61	38.15	7.47	7.11	145
C/D only :	Parking costs = 26.59 Household business vehicles = 0.186 Percent travel cost paid for = 17.58				

Source: Hensher (1986, p.14). Standard deviations given by Hensher are omitted.

Parameter Estimates (t Values in Parentheses)

(1)	0.8973	(4.86)	(2)	-2.2154	(-10.36)
(3)	1.3286	(9.10)	(4)	-0.0227	(-4.70)
(5)	-0.1336	(-6.68)	(6)	-0.0672	(-5.44)
(7)	-0.0063	(-5.03)	(8)	-0.0086	(-5.05)
(9)	0,4524	(1,83)	(10)	-0,0119	(3,71)
Log likelihood at $\beta = 0$		= -2017.1			
Log likelihood (sample shares)		= -1426.6			
Log likelihood at convergence		= -598.2			

Parameter Estimates

	C/D	C/P	Train	Bus
Probability	0.88625	0.03799	0.01390	0.06186
Predicted *N*	1290	55	20	90
Actual *N*	953	78	279	145

Elasticities of Probabilities with Respect to In-vehicle Cost

Mode	C/D	C/P	Train	Bus
CD	-0.077	0.253	0.253	0.253
CP	0.002	-0.013	0.002	0.002

Train	0.098	0.098	-0.231	0.098
Bus	0.042	0.042	0.042	-0.292

Quelle: Greene 1993

9 NESTED LOGIT und HETEROSKEDASTISCHES LOGIT Modell

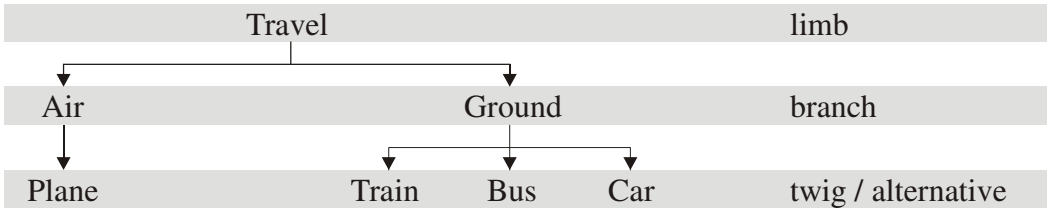
9.1 NESTED LOGIT Modell

Ist die IIA-Annahme für ein multiples Entscheidungsmodell (MNL oder CONDITIONAL LOGIT) nicht erfüllt, dann ist das mehrstufige Entscheidungsmodell, das NESTED LOGIT Modell, eine sequentielle Anwendung multipler Entscheidungsmodelle, eine Spezifikationsmöglichkeit. In bestimmten Alternativgruppen (Nestern) gilt zwar noch die IIA-Annahme, nicht mehr aber zwischen den Nestern. Das NESTED LOGIT Modell erlaubt also Entscheidungen für bestimmte Gruppen von Alternativen. Die Fehler des entsprechenden RUM–Modells sind innerhalb der Gruppen homoskedastisch (im Beispiel unten in der Gruppe Ground), nicht aber zwischen den Gruppen (Ground und Air), dort ist Heteroskedastie erlaubt.

‚Nested'-Modelle erlauben eine mehrstufige Auswahl, die bis zur Auswahl aus Alternativen auf dem untersten Level eines Entscheidungsbaumes – über einen Ast (‚limb') und dem darunter liegenden Zweig (‚branch') – führt (vgl. z.B. Cameron and Trivedi 2005, S. 507 ff, Greene 2003, S. 725 ff, Greene 1995, S. 512 ff, Börsch-Supan 1987, McFadden 1981).

Als Beispiele können alle Probleme multipler Entscheidungen (siehe die Beispiele zu Beginn dieses Kapitels) angeführt werden. Ergänzend können Gemeinsamkeiten von Alternativgruppen (Nester) untersucht werden. Innerhalb dieser Nester gilt dann wieder ein MNL bzw. CONDITIONAL LOGIT Modell.

Beispiel: Traveling



Four level example: trunk, limb, branch, twig (Stamm, Ast, Zweig, dünner Zweig)

Auswahlwahrscheinlichkeit der Alternative k (twig/alternative, z.B. ein Bus) in Zweig j (branch, z.B. Ground) und Ast i (Travel)

$$P(k \mid j,i) = \frac{\exp\left(\beta' x_{k|j,i}\right)}{\sum_{n|j,i}\exp\left(\beta' x_{n|ji}\right)} = \frac{\exp\left(\beta' x_{k|j,i}\right)}{\exp\left(I_{j|i}\right)}$$

wobei

$I_{j/i} = \ln \sum_{k=1}^{j} e^{\beta' x_{k|j}}$ der **Inklusionswert (inclusive value)** für branch j ist

Bedingte Wahrscheinlichkeit des nächsten Baum-Niveaus

$$P(j \mid i) = \frac{\exp\left(\alpha' y_{j|i} + \tau_{j|i} I_{j|i}\right)}{\sum_{n|i} \exp\left(\alpha' y_{m|i} + \tau_{m|i} I_{m|i}\right)} = \frac{\exp\left(\alpha' y_{j|i} + \tau_{j|i} I_{j|i}\right)}{\exp\left(I_i\right)}$$

wobei I_i der Inklusionswert für den Ast i ist.

Wahrscheinlichkeit für einen Ast i

$$P(i) = \frac{\exp\left(\gamma' z_i + \tau_i I_i\right)}{\sum_l \exp\left(\gamma' z_l + \tau_l I_l\right)}$$

Mit z_i als Attribute der Äste.

Die **unbedingte Auswahlwahrscheinlichkeit** ist dann

$$P(k,j,i) = P(k \mid j,i)\, P(j \mid i)\, P(i)$$

und zugleich Teil der zu maximierenden Likelihood Funktion, die die Parameter $\beta_1, \beta_2, \ ..., \ \beta_{nx}, \ \alpha_1, \ \alpha_2, \ ..., \ \alpha_{ny}, \ \gamma_1, \ \gamma_2, \ ... \ \gamma_{nz}, \ \tau_1, \tau_M, \sigma_1, ... \tau_k$ schätzt.

Für jede Auswahl, Zweig und Ast mag es unterschiedliche Nutzenfunktionen $u(Auswahl) = \beta' x + \varepsilon$ geben.

Ein **‚inclusive value'** I für eine Stufe (z.B. ein Zweig (branch)) erlaubt die Heteroskedastizität der Fehlerterme zwischen den jeweiligen Stufen (z.B. Zweigen (branches)). Wäre der Parameter τ des ‚inclusive value' gleich Eins, dann wäre die IIA-Annahme erfüllt und ein ‚normales' MNL Modell ohne Nester passend (also Homoskedastizität der Varianzen aller Alternativen). Weicht der ‚inclusive value'-Parameter τ von Eins aber ab, dann ist die IIA-Annahme zwischen den Stufen nicht erfüllt, also ein NESTED LOGIT Modell der angemessene Ansatz. Im ‚nested'-Fall ist dann der zu schätzende Parameter τ des entsprechenden ‚inclusive values' ungleich Eins.

McFadden 1978 hat gezeigt, dass auch das NESTED LOGIT Modell aus einem stochastischen Nutzenmodell abgeleitet werden kann. Das MULTINOMIALE LOGIT Modell wird unter der Annahme abgeleitet, dass die Residuen in der stochastischen Nutzenfunktion unabhängig extremwertverteilt sind. Im NESTED LOGIT Modell ist die entsprechende Annahme eine *generalisierte Extremwertverteilung* (GEV) der Residuen aller Alternativen (mehr dazu in Cameron und Trivedi, S.508, Maddala 1983, S. 70 ff).

Schätzung des NESTED LOGIT Modells

Es gibt zwei Möglichkeiten ein NESTED LOGIT Modell zu schätzen, ein ‚limited information, two step maximum likelihood' Ansatz und ein 'full information maximum likelihood' Ansatz.

Limited Information Maximum Likelihood Schätzung

Ein zweistufiger LIML Ansatz mit

1. Schätze β indem die Wahl innerhalb der Zweige (branches) als einfaches CONDITIONAL/MULTINOMIAL LOGIT Modell formuliert wird (z.B. aus der Auswahlwahrscheinlichkeit $P(k \mid j,i)$)

2. Berechne die inclusive values für alle Zweige des Modells ($I_{j/i} = \ln \sum_{k=1}^{j} e^{\beta' x_k \mid j}$) und schätze die Parameter γ und τ, indem die Wahl zwischen den Zweigen (branches) als ein CONDITIONAL oder MULTINOMIAL LOGIT Modell mit den Attributen z_i und I_i (Ast-Wahrscheinlichkeit $P(i)$) modelliert wird.

Full information Maximum Likelihood Schätzung

als FIML Ansatz mit der log-likelihood

$$\ln L = \sum_{i=1}^{n} \ln(PROB(twig \mid branch))PROB(branch))_i .$$

Der (einstufige) FIML Ansatz ist effizienter als die zweistufige Schätzung. Geschätzt werden $\beta, \gamma \text{und} \tau$. Ist der geschätzte Wert für $\tau = 1$ dann liegt ein MULTINOMIALES bzw. CONDITIONAL LOGIT Modell vor und die „Nester“ haben keine Bedeutung.

9.2 HETEROSKEDASTISCHES LOGIT (HEV) Modell

Im MULTINOMIALEN LOGIT und dem CONDITIONAL LOGIT Modell wird von der Homoskedastizität der Fehlerterme aller Alternativen ausgegangen und mit der IIA-Annahme getestet. Eine Erweiterung des CONDITIONAL LOGIT Modells, dass die evtl. Schwierigkeiten der Baumbildung des NESTED LOGIT Modells nicht hat, ist das HETEROSKEDASTISCHE LOGIT Modell (HETEROSCEDASTIC EXTREME VALUE (HEV) Modell) von Bhat 1995 und Allenby und Ginter 1995.

Ausgehend von stochastischer Nutzenstruktur wie in den LOGIT Modellen zuvor wird jetzt nicht mehr die Homoskedastizität gleicher Residuen-Varianzen ($\pi^2/6$) der Alternativen angenommen, sondern Heteroskedastizität erlaubt. Bei sonst gleicher Struktur wird also eine gestreute Varianz über die Alternativen mit $Var(\varepsilon_{ij}) = \sigma_j$ und der Normierung $\sigma_m = 1$ zugelassen.

9.3 Beispiel NESTED und HETEROSKEDASTISCHES LOGIT Modell

Wahl eines Verkehrsmittels (Greene 2003, S. 729–734)

21.7.8 APPLICATION: CONDITIONAL LOGIT MODEL FOR TRAVEL MODE CHOICE

Hensher and Greene [Greene (1995a)] report estimates of a model of travel mode choice for travel between Sydney and Melbourne, Australia. The data set contains 210 observations on choice among four travel modes, *air, train, bus,* and *car*. (See Appendix Table F21.2.) The attributes used for their example were: choice-specific constants; two choice-specific continuous measures; GC, a measure of the generalized cost of the travel that is equal to the sum of in-vehicle cost, INVC and a wagelike measure

TABLE 21.10 Summary Statistics for Travel Mode Choice Data

	GC	*TTME*	*INVC*	*INVT*	*HINC*	*Number Choosing*	*p*	*True prop.*
Air	102.648	61.010	85.522	133.710	34.548	58	0.28	0.14
	113.522	46.534	97.569	124.828	41.274			
Train	130.200	35.690	51.338	608.286	34.548	63	0.30	0.13
	106.619	28.524	37.460	532.667	23.063			
Bus	115.257	41.650	33.457	629.462	34.548	30	0.14	0.09
	108.133	25.200	33.733	618.833	29.700			
Car	94.414	0	20.995	573.205	34.548	59	0.28	0.64
	89.095	0	15.694	527.373	42.220			

Note: The upper figure is the average for all 210 observations. The lower figure is the mean for the observations that made that choice.

times INVT, the amount of time spent traveling; and TTME, the terminal time (zero for car); and for the choice between air and the other modes, HINC, the household income. A summary of the sample data is given in Table 21.10. The sample is **choice based** so as to balance it among the four choices—the true population allocation, as shown in the last column of Table 21.10, is dominated by drivers.

The model specified is

$$U_{ij} = \alpha_{air} d_{i,air} + \alpha_{train} d_{i,train} + \alpha_{bus} d_{i,bus} + \beta_G GC_{ij} + \beta_T TTME_{ij} + \gamma_H d_{i,air} HINC_i + \varepsilon_{ij}.$$

where for each j, ε_{ij} has the same independent, type 1 extreme value distribution,

$$F_\varepsilon(\varepsilon_{ij}) = \exp(-\exp(-\varepsilon_{ij}))$$

which has standard deviation $\pi^2/6$. The mean is absorbed in the constants. Estimates of the conditional logit model are shown in Table 21.11. The model was fit with and without the corrections for choice based sampling. Since the sample shares do not differ radically from the population proportions, the effect on the estimated parameters is fairly modest. Nonetheless, it is apparent that the choice based sampling is not completely innocent. A cross tabulation of the predicted versus actual outcomes is given in Table 21.12. The predictions are generated by tabulating the integer parts of $m_{jk} = \sum_{i=1}^{210} \hat{p}_{ij} d_{ik}$.

TABLE 21.11 Parameter Estimates (t Values in Parentheses)

	Unweighted Sample		*Choice Based Weighting*	
	Estimate	*t Ratio*	*Estimate*	*t Ratio*
β_G	−0.15501	−3.517	−0.01333	−2.724
β_T	−0.19612	−9.207	−0.13405	−7.164
γ_H	0.01329	1.295	−0.00108	−0.087
α_{air}	5.2074	6.684	6.5940	5.906
α_{train}	3.8690	8.731	3.6190	7.447
α_{bus}	3.1632	7.025	3.3218	5.698
Log likelihood at $\beta = 0$	−291.1218		−291.1218	
Log likelihood (sample shares)	−283.7588		−223.0578	
Log likelihood at convergence	−199.1284		−147.5896	

TABLE 21.12 Predicted Choices Based on Model Probabilities (Predictions Based on Choice Based Sampling are in Parentheses.)

	Air	*Train*	*Bus*	*Car*	*Total (Actual)*
Air	32 (30)	8 (3)	5 (3)	13 (23)	58
Train	7 (3)	37 (30)	5 (3)	14 (27)	63
Bus	3 (1)	5 (2)	15 (4)	6 (12)	30
Car	16 (5)	13 (5)	6 (3)	25 (45)	59
Total (Predicted)	58 (39)	63 (40)	30 (23)	59 (108)	210

$j, k = air, train, bus, car$, where $\hat{p}_{ij}$ is the predicted probability of outcome j for observation i and d_{ik} is the binary variable which indicates if individual i made choice k.

Are the odds ratios *train/bus* and *car/bus* really independent from the presence of the *air* alternative? To use the Hausman test, we would eliminate choice *air*, from the choice set and estimate a three-choice model. Since 58 respondents chose this mode, we would lose 58 observations. In addition, for every data vector left in the sample, the air specific constant and the interaction, $d_{i,air} \times HINC_i$ would be zero for every remaining individual. Thus, these parameters could not be estimated in the restricted model. We would drop these variables. The test would be based on the two estimators of the remaining four coefficients in the model, $[\beta_G, \beta_T, \alpha_{train}, \alpha_{bus}]$. The results for the test are as shown in Table 21.13.

The hypothesis that the odds ratios for the other three choices are independent from *air* would be rejected based on these results, as the chi-squared statistic exceeds the critical value.

Since IIA was rejected, they estimated a nested logit model of the following type:

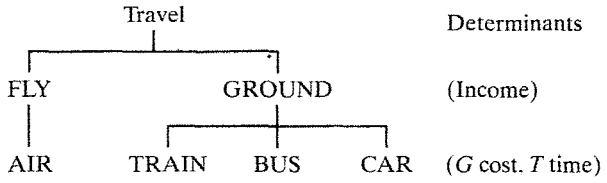


TABLE 21.13 Results for IIA Test

	Full Choice Set				*Restricted Choice Set*			
	β_G	β_T	α_{train}	α_{bus}	β_G	β_T	α_{train}	α_{bus}
Estimate	*−0.0155*	*−0.0961*	*3.869*	*3.163*	*−0.0639*	*−0.0699*	*4.464*	*3.105*
	Estimated Asymptotic Covariance Matrix				***Estimated Asymptotic Covariance Matrix***			
β_G	0.194e-5				0.000101			
β_T	−0.46e-7	0.000110			−0.0000013	0.000221		
α_{train}	−0.00060	−0.0038	0.196		−0.000244	−0.00759	0.410	
α_{bus}	*−0.00026*	*−0.0037*	*0.161*	*0.203*	−0.000113	−0.00753	0.336	0.371

Note: 0.nnne-p indicates times 10 to the negative p power.
$H = 33.3363$. Critical chi-squared[4] = 9.488.

TABLE 21.14 Estimates of a Mode Choice Model (Standard Errors in Parentheses)

Parameter	*FIML Estimate*		*LIML Estimate*		*Unconditional*	
α_{air}	6.042	(1.199)	−0.0647	(2.1485)	5.207	(0.779)
α_{bus}	4.096	(0.615)	3.105	(0.609)	3.163	(0.450)
α_{train}	5.065	(0.662)	4.464	(0.641)	3.869	(0.443)
β_{GC}	−0.03159	(0.00816)	−0.06368	(0.0100)	−0.1550	(0.00441)
β_{TTME}	−0.1126	(0.0141)	−0.0699	(0.0149)	−0.09612	(0.0104)
γ_H	0.01533	(0.00938)	0.02079	(0.01128)	0.01329	(0.0103)
τ_{flyu}	0.5860	(0.141)	0.2266	(0.296)	1.0000	(0.000)
τ_{ground}	0.3890	(0.124)	0.1587	(0.262)	1.0000	(0.000)
σ_{fly}	2.1886	(0.525)	5.675	(2.350)	1.2825	(0.000)
σ_{ground}	3.2974	(1.048)	8.081	(4.219)	1.2825	(0.000)
$\log L$	−193.6561		−115.3354 + (−87.9382)		−199.1284	

Note that one of the branches has only a single choice, so the conditional probability, $P_{j|fly} = P_{air|fly} = 1$. The model is fit by both FIML and LIML methods. Three sets of estimates are shown in Table 21.14. The set marked "unconditional" are the simple conditional (multinomial) logit (MNL) model for choice among the four alternatives that was reported earlier. Both inclusive value parameters are constrained (by construction) to equal 1.0000. The FIML estimates are obtained by maximizing the full log likelihood for the nested logit model. In this model,

$$\text{Prob}(choice \mid branch) = P(\alpha_{air}d_{air} + \alpha_{train}d_{train} + \alpha_{bus}d_{bus} + \beta_G GC + \beta_T TTME),$$

$$\text{Prob}(branch) = P(\gamma d_{air} HINC + \tau_{fly} IV_{fly} + \tau_{ground} IV_{ground}),$$

$$\text{Prob}(choice, branch) = \text{Prob}(choice \mid branch) \times \text{Prob}(branch).$$

Finally, the limited information estimator is estimated in two steps. At the first step, a choice model is estimated for the three choices in the ground branch:

$$\text{Prob}(choice \mid ground) = P(\alpha_{train}d_{train} + \alpha_{bus}d_{bus} + \beta_G GC + \beta_T TTME)$$

This model uses only the observations that chose one of the three ground modes; for these data, this subset was 152 of the 210 observations. Using the estimates from this model, we compute, for all 210 observations, $IV_{fly} = \log[\exp(\mathbf{z}'_{air}\boldsymbol{\beta})]$ for *air* and 0 for ground, and $IV_{ground} = \log[\sum_{j=ground} \exp(\mathbf{z}'_j\boldsymbol{\beta})]$ for ground modes and 0 for *air*. Then, the choice model

$$\text{Prob}(branch) = P(\alpha_{air}d_{air} + \gamma_H d_{air} HINC + \tau_{fly} IV_{fly} + \tau_{ground} IV_{ground})$$

is fit separately. Since the Hessian is not block diagonal, the FIML estimator is more efficient. To obtain appropriate standard errors, we must make the Murphy and Topel correction for two-step estimation; see Section 17.7 and Theorem 17.8. It is simplified a bit here because different samples are used for the two steps. As such, the matrix $\mathbf{R}$ in the theorem is not computed. To compute $\mathbf{C}$, we require the matrix of derivatives of log Prob(*branch*) with respect to the direct parameters, α_{air}, γ_H, τ_{fly}, τ_{ground}, and with respect to the choice parameters, $\boldsymbol{\beta}$. Since this model is a simple binomial (two choice) logit model, these are easy to compute, using (21-19). Then the corrected asymptotic covariance matrix is computed using Theorem 17.8 with $\mathbf{R} = \mathbf{0}$.

TABLE 21.15 Estimates of a Heteroscedastic Extreme Value Model (Standard Errors in Parentheses)

Parameter	*HEV Estimate*		*Nested Logit Estimate*		*Restricted HEV*	
α_{air}	7.8326	(10.951)	6.062	(1.199)	2.973	(0.995)
α_{bus}	7.1718	(9.135)	4.096	(0.615)	4.050	(0.494)
α_{train}	6.8655	(8.829)	5.065	(0.662)	3.042	(0.429)
β_{GC}	−0.05156	(0.0694)	−0.03159	(0.00816)	−0.0289	(0.00580)
β_{TTME}	−0.1968	(0.288)	−0.1126	(0.0141)	−0.0828	(0.00576)
γ	0.04024	(0.0607)	0.01533	(0.00938)	0.0238	(0.0186)
τ_{fly}		—	0.5860	(0.141)		—
τ_{ground}		—	0.3890	(0.124)		—
θ_{air}	0.2485	(0.369)			0.4959	(0.124)
θ_{train}	0.2595	(0.418)			1.0000	(0.000)
θ_{bus}	0.6065	(1.040)			1.0000	(0.000)
θ_{car}	1.0000	(0.000)			1.0000	(0.000)
Implied Standard Deviations						
σ_{air}	5.161	(7.667)				
σ_{train}	4.942	(7.978)				
σ_{bus}	2.115	(3.623)				
σ_{car}	1.283	(0.000)				
ln L	−195.6605		−193.6561		−200.3791	

The likelihood ratio statistic for the nesting (heteroscedasticity) against the null hypothesis of homoscedasticity is $-2[-199.1284-(-193.6561)] = 10.945$. The 95 percent critical value from the chi-squared distribution with two degrees of freedom is 5.99, so the hypothesis is rejected. We can also carry out a Wald test. The asymptotic covariance matrix for the two inclusive value parameters is [0.01977/0.009621, 0.01529]. The Wald statistic for the joint test of the hypothesis that $\tau_{fly} = \tau_{ground} = 1$, is

$$W = (0.586 - 1.0 \quad 0.389 - 1.0) \begin{bmatrix} 0.1977 & 0.009621 \\ 0.009621 & 0.01529 \end{bmatrix}^{-1} \begin{pmatrix} 0.586 - 1.0 \\ 0.389 - 1.0 \end{pmatrix} = 24.475$$

The hypothesis is rejected, once again.

The nested logit model was reestimated under assumptions of the heteroscedastic extreme value model. The results are shown in Table 21.15. This model is less restrictive than the nested logit model. To make them comparable, we note that we found that $\sigma_{air} = \pi/(\tau_{air}\sqrt{6}) = 2.1886$ and $\sigma_{train} = \sigma_{bus} = \sigma_{car} = \pi/(\tau_{ground}\sqrt{6}) = 3.2974$. The heteroscedastic extreme value (HEV) model thus relaxes one variance restriction, because it has three free variance parameters instead of two. On the other hand, the important degree of freedom here is that the HEV model does not impose the IIA assumption anywhere in the choice set, whereas the nested logit does, within each branch.

A primary virtue of the HEV model, the nested logit model, and other alternative models is that they relax the IIA assumption. This assumption has implications for the cross elasticities between attributes in the different probabilities. Table 21.16 lists the estimated elasticities of the estimated probabilities with respect to changes in the generalized cost variable. Elasticities are computed by averaging the individual sample values rather than computing them once at the sample means. The implication of the IIA

TABLE 21.16 Estimated Elasticities with Respect to Generalized Cost

	Cost Is That of Alternative			
Effect on	***Air***	***Train***	***Bus***	***Car***
Multinomial Logit				
Air	−1.136	0.498	0.238	0.418
Train	0.456	−1.520	0.238	0.418
Bus	0.456	0.498	−1.549	0.418
Car	0.456	0.498	0.238	−1.061
Nested Logit				
Air	−0.858	0.332	0.179	0.308
Train	0.314	−4.075	0.887	1.657
Bus	0.314	1.595	−4.132	1.657
Car	0.314	1.595	0.887	−2.498
Heteroscedastic Extreme Value				
Air	−1.040	0.367	0.221	0.441
Train	0.272	−1.495	0.250	0.553
Bus	0.688	0.858	−6.562	3.384
Car	0.690	0.930	1.254	−2.717

assumption can be seen in the table entries. Thus, in the estimates for the multinomial logit (MNL) model, the cross elasticities for each attribute are all equal. In the nested logit model, the IIA property only holds within the branch. Thus, in the first column, the effect of GC of air affects all ground modes equally, whereas the effect of GC for train is the same for bus and car but different from these two for air. All these elasticities vary freely in the HEV model.

Quelle: Greene 2003, S. 729–734

10 Modelle für geordnete Kategorien: 'Ordered response models'

10.1 LOGIT- und PROBIT-Modelle für geordnete Kategorien

Einige multinominale Auswahlvariablen können geordnet sein, also eine ordinale Rangfolge aufweisen (Greene 2008b, Greene and Hensher 2010, Greene 1997, chapt. 19-8, Maddala 1983, chapt. 2.13).

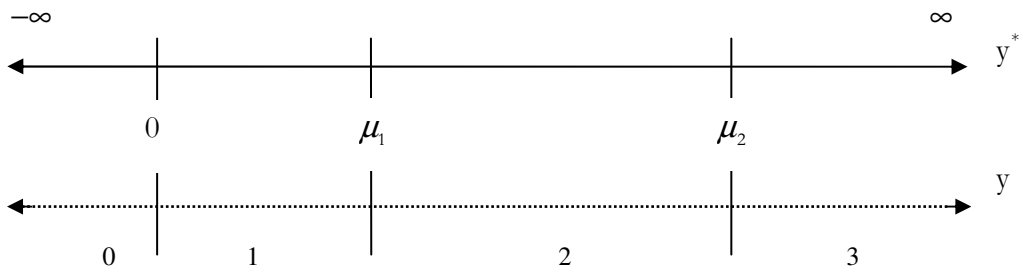
Beispiele: Happiness/Zufriedenheit scores, Resultat eines Tests: gut, schlecht, miserabel; Ausbildungsstufen; Versicherungsabdeckung: nichts, Teilkasko, Vollkasko; Beschäftigung: arbeitslos, Teilzeit, Vollzeit; Meinungsumfragen....

Ordered response model (analog der binomialen Vorgehensweise) für J+1 Kategorien

$$y^* = \beta' x + \varepsilon \qquad y^* \text{ unbeobachtet, latent}$$

$$\begin{aligned} y &= 0 \quad falls \quad y^* \le 0 \\ &= 1 \quad 0 < y^* \le \mu_1 \\ &= 2 \quad \mu_1 < y^* \le \mu_2 \\ &= \ldots \\ &= J \quad \mu_{J-1} \le y^* \end{aligned}$$

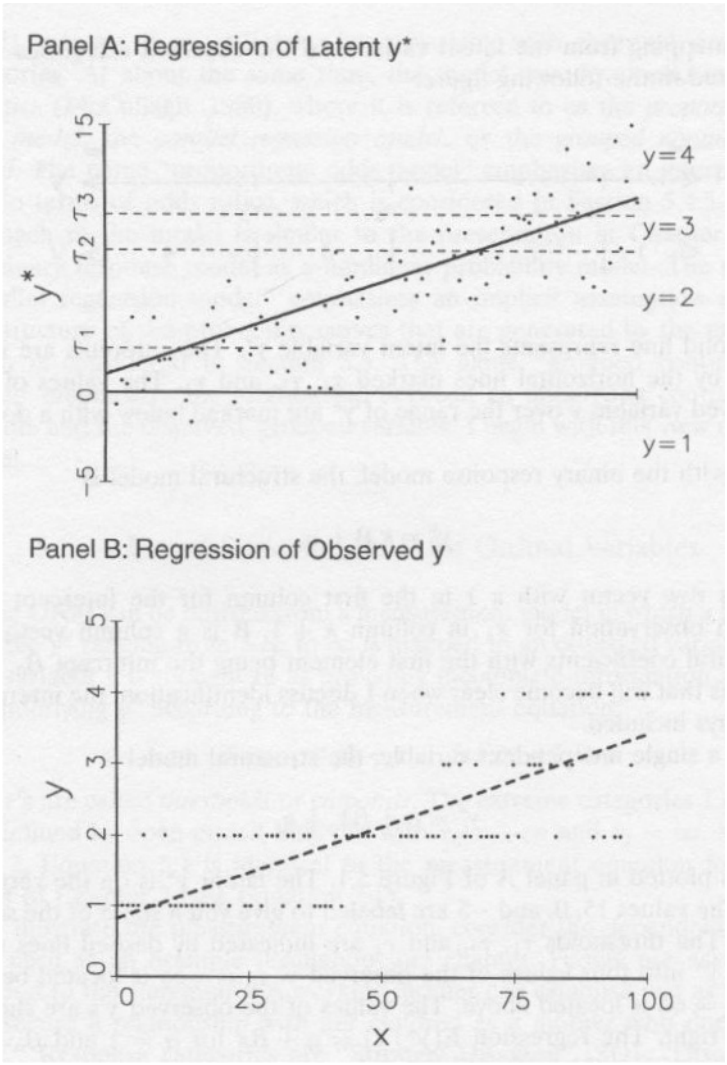
Zu schätzen sind die unbekannten Schwellenwerte μ und die Koeffizienten β.



Ist ε normalverteilt → ORDERED PROBIT

ε logistisch verteilt → ORDERED LOGIT

Das behandelte binomiale PROBIT- oder LOGIT-Modell ist mit J = 1 ein Sonderfall. Damit ist es eine Erweiterung des multinomialen Ansatzes.

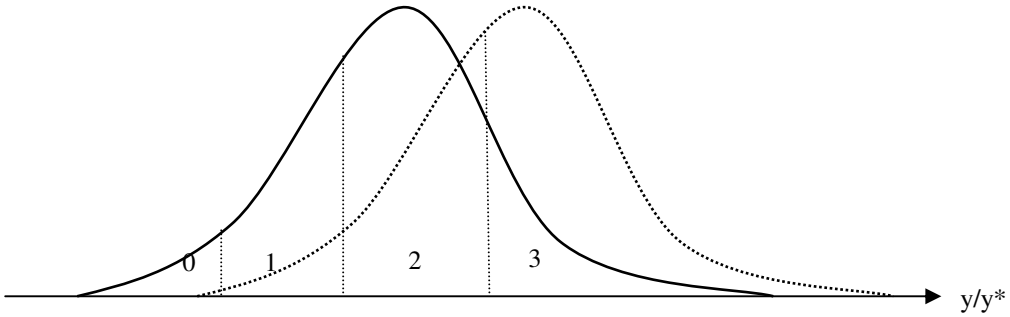


Latente und beobachtete Probanden im OR-Modell (Long und Freese 1997, S. 118)

ORDERED PROBIT

$$P(y=0) = P(\beta'x + \varepsilon < 0) = P(\varepsilon < -\beta'x)$$
$$= \Phi(-\beta'x)$$
$$P(y=1) = \Phi(\mu_1 - \beta'x) - \Phi(-\beta'x)$$
$$P(y=2) = \Phi(\mu_2 - \beta'x) - \Phi(\mu_1 - \beta'x)$$
.....
$$P(y=J) = 1 - \Phi(\mu_{J-1} - \beta'x)$$
und $0 < \mu_1 < \mu_2 < \ldots < \mu_{J-1}$

Wie sich eine Veränderung der latenten Variable y^* auf die Wahrscheinlichkeit einer bestimmten Kategorie j anzugehören auswirkt, lässt sich durch eine Verschiebung der Dichtefunktion erkennen (siehe folgende Abbildung).



Diese Generalisierung des binomialen PROBIT-Modells wird entsprechend über den Maximum Likelihood Ansatz geschätzt.

$$\text{Sei } d_{ij} \begin{cases} = 1 & falls\ y_i\ in\ die\ Kategorie\ j\ fällt \\ = 0 & sonst \end{cases}$$

$$P(d_{ij}=1) = \Phi\left(\mu_j - \beta' x_i\right) - \Phi\left(\mu_{j-1} - \beta' x_i\right)$$

$$L = \prod_i \prod_j \left[\Phi\left(\mu_j - \beta' x_i\right) - \Phi\left(\mu_{j-1} - \beta' x_i\right)\right]^{d_{ij}}$$

Mit Newton-Raphson bspw. werden dann die μ s und β geschätzt. (Pratt 1981 hat gezeigt, dass dieses Modell ein globales Maximum hat).

Für die Berechnung dann der alternativenspezifischen Wahrscheinlichkeiten sind die μs mit einzubeziehen: $P(y=j) = \Phi\left(\mu_j - \beta' x\right) - \Phi\left(\mu_{j-1} - \beta' x\right)$.

Die Inverse der Informationsmatrix (Erwartungswert der Hesse'sche Matrix der 2. Ableitungen im Optimum) ergibt die asymptotische Kovarianzmatrix von μ und β und mit deren Varianz dann t- und andere Testwerte zur Überprüfung der Signifikanz der Modellergebnisse berechnet werden können.

Marginale Effekte

$$\frac{\partial P(y=0)}{\partial x} = -\phi\left(\beta' x\right)\beta$$

$$\frac{\partial P(y=1)}{\partial x} = \left[\phi\left(-\beta\ x\right) - \phi\left(\mu - \beta' x\right)\right]\beta$$

$$\frac{\partial P(y=2)}{\partial x} = \phi\left(\mu - \beta' x\right)\beta$$

Das ORDERED LOGIT Modell wird analog dann mit der logistischen Verteilungsfunktion $\Lambda(.)$ anstelle der Normalverteilung $\Phi(.)$ gebildet.

Ein dem ORDERED RESPONSE Modell verwandtes polytomes Modell ist das SEQUENTIAL RESPONSE Modell (Maddala 1983, S. 49f), das hintereinander wiederholt dichotome Entscheidungen fällt.

10.2 GENERALIZED ORDERED PROBIT (GOP) Modell

Das ORDERED PROBIT Modell trifft implizit die sogenannte *parallel regression assumption* (PRA) (Long 1997, Greene und Hensher 2010). Sie besagt, dass die Wirkungsweise der erklärenden Variablen für alle Kategorien identisch ist. Das bedeutet, dass getrennte PROBIT Modelle für jede Kategorie die gleichen geschätzten Koeffizienten erbringt.

Long und Freese 2006 schlagen zwei Testverfahren zur Überprüfung der PRA vor: einen approximativen Likelihood-Ratio Test und den Brant Test (Brant 1990). Neben dem Gesamtansatz, kann für jede Einflussgröße mit dem Brant Test die Nullhypothese, die kategoriespezifischen Koeffizienten sind alle gleich, überprüft werden. Hohe p-value Werte sind ein Indiz dafür, dass die PRA nicht verletzt ist. In Stata steht der Brant Test nach dem ORDERED LOGIT Modell zur Verfügung.

Generalisierte Modelle für ordinal abhängige Variablen, die eine Lockerung der PRA ermöglichen, werden von Winkelmann und Bös 2006 beschrieben. Das GENERALIZED ORDERED PROBIT Modell

$$
\begin{aligned}
P(y=0) &= P(\beta' x + \varepsilon < 0) = P(\varepsilon < -\beta' x) \\
&= \Phi\left(-\beta_0' x\right) \\
P(y=1) &= \Phi\left(\mu_1 - \beta_1' x\right) - \Phi\left(-\beta_0' x\right) \\
P(y=2) &= \Phi\left(\mu_2 - \beta_2' x\right) - \Phi\left(\mu_1 - \beta_1' x\right) \\
&\ldots\ldots \\
P(y=J) &= 1 - \Phi\left(\mu_{J-1} - \beta_{J-1}' x\right) \\
\textit{und} \quad & 0 < \mu_1 < \mu_2 < \ldots < \mu_{J-1}
\end{aligned}
$$

hat also für jede Kategorie einen eigenen geschätzten Parametervektor β_j und berücksichtigt dabei aber die Ordungsstruktur auch noch über die μ s.

Die Schätzung erfolgt analog zum ORDERED PROBIT Modell mit einem Maximum Likelihood Schätzer.

Für die Berechnung dann der alternativenspezifischen Wahrscheinlichkeiten sind die μ s mit einzubeziehen: $P(y=j) = \Phi\left(\mu_j - \beta_j' x\right) - \Phi\left(\mu_{j-1} - \beta_{j-1}' x\right)$.

10.3 Beispiele 'Ordered response models'

- Zum Einkommen der Freien Berufe – Eine Ordered Probit-Analyse ihrer Determinanten auf Basis der FFB-Onlineumfrage (Joachim Merz und Peter Paic (2005), Zum Einkommen der Freien Berufe – Eine Ordered Probit-Analyse ihrer Determinanten auf Basis der FFB-Onlineumfrage, in: Schulte, R. (Hg.), Ergebnisse der MittelstandsForschung, Lit-Verlag, Münster – London, 327-348.)
- Arbeitszufriedenheit und Flexibilität – Eine GENERALIZED ORDERED PROBIT Analyse (Hanglberger 2011)
- Job assignment of new Navy recruits: ORDERED PROBIT (Greene 1993, S. 675) Tab. 5.2
- Congressional voting determinants on the 1965 Medicare Bill (McKelvey and Zavoina 1975)
- OLOGIT, OPROBIT und GOLOGIT und GOPROBIT für das MULTINOMIALE LOGIT Modell Beispiel Berufliche Stellung (Abschnitt 8.1)

10.3.1 Zum Einkommen der Freien Berufe – Eine Ordered Probit-Analyse ihrer Determinanten auf Basis der FFB-Onlineumfrage (Merz und Paic 2005)

…

5 Mikroökonometrisches Modell: Ordered Probit

Kommen wir nun zur multivariaten Analyse, um die Signifikanz der miteinander konkurrierenden Erklärungsfaktoren der einzelnen Hypothesen untersuchen zu können.

Für eine höhere Antwortbereitschaft haben wir die Höhe des freiberuflichen Einkommens im FFB-Onlinefragebogen in insgesamt zehn Klassen erfasst. Damit ist eine ordinale Skala gegeben, für die sich als Schätzmodell ein Ordered Probit Ansatz empfiehlt (vgl. Greene 1997, S. 875 ff; Moosmüller 2004, 203 ff). Ein Ordered Probit Ansatz bietet zudem gegenüber einem Schätzansatz mit Klassenmittelwerten den Vorteil, dass die Klassenmitte der nach oben offenen höchsten Einkommensklasse – da beliebig – nicht zu spezifizieren ist.

Das Ordered Probit Modelle geht bekanntlich von einem latenten Modell aus mit

(7)
$$y_i^* = \beta' x_i + \varepsilon_i \qquad \varepsilon_i \; verteilt\, mit\, N(0,1)$$

Der Vektor x_i enthält die Merkmalsausprägungen der erklärenden Variablen, β ist ein Vektor mit den unbekannten Parametern des Modells und ε_i ist ein standardnormalverteilter Störterm. Die latente (Einkommens-) Variable wird durch die Klassengrenzen eingegrenzt mit

(8)
$$y_i = \begin{cases} 0, & falls \quad -\infty < y_i^* \le \mu_0 \\ 1, & falls \quad \mu_0 < y_i^* \le \mu_1 \\ & \vdots \\ J, & falls \quad \mu_{J-1} < y_i^* \le \infty \end{cases}$$

Prinzipiell werden im Ordered Probit Modell die Klassengrenzen mit den nur bekannten diskreten Klassennummern geschätzt. Diese sind mit den Antwortklassen aber bereits gegeben. Mit bekannten Klassengrenzen ist es möglich, die Varianz von ε zu schätzen, die sonst im Ordered Probit Modell angenommen werden muss. Für die Bestimmung der Wahrscheinlichkeit, bei gegebenen erklärenden Größen zu einer bestimmten Einkommensklasse zu gehören, werden die weiteren Schätzergebnisse des Ordered Probit Modells bezüglich β benötigt:

(9)
$$P(y_i = j) = \begin{cases} P(\mu_{j-1} < y_i^* \le \mu_j) \\ P(\mu_{j-1} < \beta' x_i + \varepsilon_i \le \mu_j) \\ P(\mu_{j-1} - \beta' x_i \le \varepsilon_i < \mu_j - \beta' x_i) \\ \Phi(\mu_j - \beta' x_i) - \Phi(\mu_{j-1} - \beta' x_i) \end{cases}$$

mit der Standardnormalverteilung Φ

In unserer Anwendung werden die 10 ordinalskalierten Einkommensklassen zu drei Einkommensklassen für eine ausreichende Besetzung der Klassen zusammengefasst. Die erste Klasse setzt sich aus dem jährlichen Bruttoeinkommen aus freiberuflicher Tätigkeit bis 37.500 € zusammen. Die zweite Klasse geht bis 74.500 €. Die dritte Klasse ist nach oben offen. Diese Klassengrenzen folgen aus der Zusammenfassung der ersten und zweiten drei und der letzten vier Klassen der zehn Originalklassen.

Ein positives Vorzeichen eines mit der Maximum Likelihood Methode geschätzten Koeffizienten gibt an, zu einer höheren Einkommensklasse zu gehören; ein negatives Vorzeichen entsprechend zu einer niedrigeren Einkommensklasse zu gehören.

6 Determinanten freiberuflichen Einkommens – Ergebnisse der multivariaten Analyse

Untersucht werden die folgend zugeordneten Hypothesen/Determinanten: Aus dem humankapitaltheoretischen Ansatz: Schulbildung in Jahren, Hochschulabschluss (Dichotom, Referenz kein Hochschulabschluss), Berufserfahrung in der derzeitigen freiberuflichen Tätigkeit in Jahren, vorherige Branchenerfahrung in Jahren, vorherige Selbständigkeit (Dichotom, Referenz keine Selbständigkeitserfahrung). Aus dem institutionellen und betrieblichen Umfeld sind es die Determinanten der Berufsgruppe: Freie Heilberufe, Freie rechts- und wirtschaftsberatende Berufe und die Freien künstlerischen, pädagogischen und publizistischen Berufe (Referenz Freie technische und naturwissenschaftliche Berufe), Kammerberuf (Dichotom, Referenz kein Kammerberuf), derivative Gründungsform (Dichotom, Referenz originäre Gründungsform), sowie allgemeine betriebliche und demografische Determinanten: Geschlecht weiblich (Dichotom, Referenz männlich), Anzahl der sozialversicherten Mitarbeiter jetzt (2003), Anzahl der sozialversicherten Mitarbeiter zum Gründungszeitpunkt, Region der Wirkungsstätte Ostdeutschland (Dichotom, Referenz Westdeutschland).

Zum Vergleich der Ergebnisse werden in Tabelle 3 dem Ordered Probit Modell zwei weitere Schätzmodelle gegenübergestellt, die allerdings die vorhandene Ordnungsstruktur nicht berücksichtigen, von daher also weniger effizient sind. Dies ist zum einen ein Multinomiales Logit (MNL) Modell, das auf einem stochastischen Nutzenkonzept für die Auswahl gleichrangiger Alternativen basiert. Zum anderen ist es der einfache klassische OLS-Schätzer, der die Klassenmitten mit einer angenommenen (arbiträren) Klassenmitte der obersten Klasse auskommen muss. Wir argumentieren, dass die Ergebnisse besonders robust sind, wenn sie - außer dem bevorzugten Ordered Probit Modell - auch von den anderen Schätzmodellen getragen werden. Die untere Einkommensklasse dient dem MNL-Modell als Referenz. Für alle Schätzverfahren werden die jeweiligen geschätzten Koeffizienten (C) und ihr Probability-Value (P) angegeben.

Quelle: Merz und Paic 2005

LIMDEP: ORDERED PROBIT Schätzung der Determinanten der Einkommen Freier Berufe

y-Variable: BRUTTO3G (unteres (0), mittleres (1) und höheres Einkommen (2))

```
--> namelist;x=ONE,SCHOOL,EXPER,EXPER2,ERFAJAHR,frau,HOCHSCHU,HEILBERU,
    RECHTBER,TECHNISC,grndori,ANZMIT,ANZMITGR,KAMMERBE,ost,erselb$

--> ORDERED;Lhs=BRUTTO3G;Rhs=x$

Normal exit from iterations. Exit status=0.
+---------------------------------------------+
| Ordered Probability Model                   |
| Maximum Likelihood Estimates                |
| Model estimated: Nov 24, 2004 at 09:49:37PM.|
| Dependent variable            BRUTTO3G      |
| Weighting variable                 None     |
```

```
| Number of observations              123     |
| Iterations completed                 23     |
| Log likelihood function       -99.73201     |
| Restricted log likelihood     -130.0080     |
| Chi squared                    60.55195     |
| Degrees of freedom                   15     |
| Prob[ChiSqd > value] =         .0000000     |
| Underlying probabilities based on Normal    |
|    Cell frequencies for outcomes            |
|  Y Count Freq  Y Count Freq  Y Count Freq   |
|  0    48 .390  1    50 .406  2    25 .203   |
+---------------------------------------------+
```

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]	Mean of X
	Index function for probability				
Constant	.5175291841	1.2717847	.407	.6841	
SCHOOL	-.1301817213	.10329590	-1.260	.2076	12.284553
EXPER	.1290063274	.48557696E-01	2.657	.0079	9.4227642
EXPER2	-.4097909596E-01	.16381401E-01	-2.502	.0124	17.652846
ERFAJAHR	.2679953703E-01	.23034339E-01	1.163	.2446	5.4390244
FRAU	-.8619101946	.27127302	-3.177	.0015	.38211382
HOCHSCHU	.3605899295	.28316557	1.273	.2029	.56910569
HEILBERU	.7759540271	.35319296	2.197	.0280	.54471545
RECHTBER	.3368001103	.43848672	.768	.4424	.15447154
TECHNISC	-.3469505724	.54752066	-.634	.5263	.97560976E-01
GRNDORI	.2592014707	.28813940	.900	.3683	.75609756
ANZMIT	.1590620069E-01	.49983352E-01	.318	.7503	2.1707317
ANZMITGR	.7890359101E-01	.53840570E-01	1.466	.1428	1.4878049
KAMMERBE	.8604543490	.34032438	2.528	.0115	.21951220
OST	-.5829655527	.47104792	-1.238	.2159	.81300813E-01
ERSELB	.1623768796	.39909771	.407	.6841	.89430894E-01
	Threshold parameters for index				
Mu(1)	1.548064593	.19064151	8.120	.0000	

(Note: E+nn or E-nn means multiply by 10 to + or -nn power.)

Cross tabulation of predictions. Row is actual, column is predicted.
Model = Probit . Prediction is number of the most probable cell.

Actual	Row Sum	0	1	2	3	4	5	6	7	8	9
0	48	36	11	1							
1	50	14	32	4							
2	25	0	14	11							
Col Sum	123	50	57	16	0	0	0	0	0	0	0

Tabelle 3: Determinanten der Einkommen Freier Berufe – Ergebnisse des Ordered Probit Modells, des Multinomialen Logit Modells und des OLS Schätzers

	Ordered Probit		OLS		MNL mittlere Einkommen		MNL höhere Einkommen	
	Coef.	P-value	Coef.	P-value	Coef.	P-value	Coef.	P-value
Konstante	.5175	.6841	10,833	.0001	1,580	.6134	2,174	.6437
Humankapital								
Schuljahre	-.1301	.2076	-.6813	.1934	-.3178	.2238	-.6595	.1015
Hochschulabschluss	.3605	.2029	.1562	.2603	-.2422	.7230	**2,441**	**.0320**
Berufserfahrung	**.1290**	**.0079**	**.5699**	**.0093**	**.2477**	**.0283**	**.3955**	**.0131**
Berufserfahrung quadriert	**-.4097**	**.0124**	**-.1772**	**.0120**	**-.7562**	**.0336**	**-.1302**	**.0160**
vorh. Branchenerfahrung	.2679	.2446	.9813	.3906	-.5506	.3930	.0607	.3715
vorh. Selbständigkeit	.1623	.6841	.8576	.6618	-.6626	.5826	1,156	.3600
Geschlecht								
Frau	**-.8619**	**.0015**	**-.3966**	**.0025**	**-.2498**	**.0125**	**-3.248**	**.0049**
Berufsgruppe								
Freie Heilberufe	**.7759**	**.0280**	**.3608**	**.0284**	**1.859**	**.0268**	1.773	.1669
Freie Rechts- u. Wirt.	.3368	.4424	.1000	.6265	-.4234	.6995	.9896	.4758
Freie Künstler etc.	-.3469	.5263	-.1242	.6218	1.675	.2438	-.3036	.8713
Gründungsform								
Derivativ	.2592	.3683	.1289	.3503	**1.533**	**.0376**	.8077	.3919
Soz-vers. Mitarbeiter								
Anzahl der Mitarbeiter	.1590	.7503	.1865	.4669	**.5321**	**.0156**	.2861	.1536
Anz. Mitarb. Gründung	.7890	.1428	.2792	.2912	-.3518	.1076	-.3214	.8538
Kammerberufe								
Kammerberuf	**.8604**	**.0115**	**.3832**	**.0255**	.1140	.9037	**2.532**	**.0135**
Region								
Ost	-.5829	.2159	-.2374	.2932	-.2408	.8261	**-2.928**	**.0948**
Modellgüte								
Log Likelihood	-99.732		----------		-79.051			
Chi-Squared	60.552***		----------		101.914***			
R-Squared	----------		.30896***		----------			
Adjust. R-Squared	----------		.39392***		----------			
n	123		123		123			

Einkommen: Jährliches Bruttoeinkommen aus freiberuflicher Tätigkeit

Quelle: FFB-Onlineerhebung Freie Berufe 2003, Merz und Paic 2005.

10.3.2 Arbeitszufriedenheit und Flexibilität – Eine GENERALIZED ORDERED PROBIT Analyse (Hanglberger 2011)

Für 31 europäische Länder (European Working Condition Survey, EWCS) hat Hanglberger mit seiner Dissertation : Arbeitszufriedenheit und Flexibilität – Internationaler Vergleich und Adaptions- und Antizipationseffekte unter anderem einen Vergleich zwischen dem ORDERED PROBIT und dem GENERALIZED ORDERED PROBIT Modell vorgenommen (Hanglberger 2011, 107):

Tabelle 3-9: GOP und ordered-probit Schätzung der Arbeitszufriedenheit im Vergleich

Merkmal	Ordered-probit	Generalisiertes ordered-probit (GOP) (β_1)	(β_2)	(β_3)
ln(Lohnsatz) aus Haupttätigkeit	**0,128***	**0,070****	**0,121***	**0,150***
	(6,63)	**(2,62)**	**(5,69)**	**(7,15)**
subjektiv gut bezahlt	**0,555***	**0,461***	**0,607***	**0,538***
	(27,95)	**(7,60)**	**(20,46)**	**(22,46)**
1./2. Dezil	**0,013**	**-0,147****	**-0,019**	**0,087***
	(0,46)	**(-2,96)**	**(-0,58)**	**(2,50)**
Stunden/Woche	-0,003	-0,001	-0,001	-0,001
	(-0,65)	(-0,30)	(-0,30)	(-0,30)
(Stunden/Woche)²/100	0,004	0,002	0,002	0,002
	(0,75)	(0,36)	(0,36)	(0,36)
Flexibilität für Arbeitgeber	-0,051*	-0,049	-0,049	-0,049
	(-1,97)	(-1,90)	(-1,90)	(-1,90)
Flexibilität für Arbeitnehmer	0,143***	0,143***	0,143***	0,143***
	(6,81)	(6,80)	(6,80)	(6,80)
nicht normale Arbeitszeiten	-0,046*	-0,047*	-0,047*	-0,047*
	(-2,19)	(-2,25)	(-2,25)	(-2,25)
Konflikt berufl./priv. Verpflichtungen	**-0,463***	**-0,368***	**-0,526***	**-0,403***
	(-19,15)	**(-8,39)**	**(-18,29)**	**(-10,99)**
körperliche Belastung mittel	-0,194***	-0,196***	-0,196***	-0,196***
	(-8,88)	(-8,94)	(-8,94)	(-8,94)
körperliche Belastung hoch	-0,365***	-0,370***	-0,370***	-0,370***
	(-14,12)	(-14,29)	(-14,29)	(-14,29)
Stress	**-0,296***	**-0,370***	**-0,338***	**-0,240***
	(-10,25)	**(-6,00)**	**(-8,72)**	**(-6,43)**
Autonomie	0,077***	0,077***	0,077***	0,077***
	(3,80)	(3,78)	(3,78)	(3,78)
Vorgesetzter	0,127***	0,127***	0,127***	0,127***
	(4,62)	(4,62)	(4,62)	(4,62)
Anspruch	0,018	0,019	0,019	0,019
	(0,85)	(0,89)	(0,89)	(0,89)
Sinn	**0,594***	**0,594***	**0,662***	**0,362***
	(18,06)	**(12,17)**	**(17,71)**	**(6,47)**
Aufstiegsmöglichkeiten	0,392***	0,398***	0,398***	0,398***
	(17,86)	(18,01)	(18,01)	(18,01)
befristet beschäftigt	**0,026**	**-0,038**	**-0,035**	**0,091****
	(1,09)	**(-0,77)**	**(-1,12)**	**(2,99)**
Arbeitsplatz unsicher	**-0,272***	**-0,335***	**-0,299***	**-0,194***
	(-10,95)	**(-7,21)**	**(-9,80)**	**(-5,38)**
Teilzeit 5 bis 29 h/Woche	-0,131*	-0,128*	-0,128*	-0,128*
	(-2,44)	(-2,45)	(-2,45)	(-2,45)
Betriebszugehörigkeit (Jahre)	-0,003	-0,003	-0,003	-0,003
	(-0,75)	(-0,83)	(-0,83)	(-0,83)
Betriebszugehörigkeit²/100	0,002	0,003	0,003	0,003
	(0,19)	(0,26)	(0,26)	(0,26)
mehrere Jobs	**0,013**	**-0,041**	**-0,069**	**0,083***
	(0,38)	**(-0,58)**	**(-1,62)**	**(2,01)**
Arbeitsweg: h/Tag	**-0,156***	**0,008**	**-0,091**	**-0,263***
	(-4,05)	**(0,11)**	**(-1,92)**	**(-5,59)**
Arbeitsweg²	**0,037****	**-0,034**	**0,008**	**0,085***
	(2,59)	**(-1,59)**	**(0,51)**	**(5,19)**
Beziehungen zu Vorgesetzten und Kollegen	**0,312***	**0,384***	**0,321***	**0,292***
	(16,33)	**(9,01)**	**(12,72)**	**(11,31)**
Branche	✓		✓	
Tätigkeit	✓		✓	
Person und Haushalt	✓		✓	
X^2 (Wald-Test)	5477,5		5691,1	
p-value (Wald-Test)	0,000		0,000	
N	17.929		17.929	
LR-Test ordered-probit vs. GOP χ^2 (26)		160,39		
p-value (LR-Test)		0,000		

Hinweise: Die Spalten β_1, β_2 und β_3 geben die drei Koeffizientenvektoren des GOP-Modells wieder; Koeffizienten der Merkmale, für die die PRA im GOP gelockert wurde fett; Koeffizienten nicht standardisiert, t-Statistiken in Klammern, * $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$; EWCS 31; Alter: 15 bis 64; mindestens fünf Wochenstunden. Eine Ergebnistabelle die auch alle Kontrollvariablen umfasst findet sich in Anhang A 6.

Quelle: Eigene Berechnungen basierend auf EWCS 2005.

10.3.3 Job assignment of new Navy recruits: ORDERED PROBIT (Quelle: Greene 2003, p. 739)

Marcus and Greene (1985) estimated an ordered probit model for the job assignments of new Navy recruits. The Navy attempts to direct recruits into job classifications in which they will be most productive. The broad classifications the authors analyzed were technical jobs with three clearly ranked skill ratings: 'medium skilled,' 'highly skilled' and 'nuclear qualified/highly skilled'. Since the assignment is partly based on the Navy's own assessment and needs and partly on factors specific to the individual, an ordered probit model was used with the following determinants.

Variablen:	y = medium skilled, highly skilled, nuclear qualified/ highly skilled
ENSPA:	dummy variable indicating that the individual entered the Navy with an A school guarantee
EDMA	educational level of entrants mother
AFQT	score on the Air Force Qualifying Test
EDYRS	completed years of education
MARR	dummy variable indicating if entrant is married
AGEAT	age at time of enlistment

Variable	Estimate	T-ratio	Mean of Var.
Constant	-4,340		
ENSPA	0,057	1,700	0,660
EDMA	0,007	0,800	12,100
AFQT	0,039	39,900	71,200
EDYRS	0,190	8,700	12,100
MARR	-0,480	-9,000	0,080
AGEAT	0,0015	0,100	18,800
μ	1,790	80,800	

	$-\hat{\beta}'x$	$\hat{\mu}-\hat{\beta}'x$	Prob (Y=0)	Prob (Y=1)	Prob (Y=2)
MARR = 0	-0,8863	0,9037	0,187	0,629	0,184
MARR = 1	-0,4063	1,3837	0,342	0,574	0,084
Change			0,155	-0,055	-0,100

Die Grenzen: 0 mit $\Phi(-\hat{\beta}'x)$ und μ mit $\Phi(\hat{\mu}-\hat{\beta}'x)$ ergibt drei Regionen.

10.3.4 OLOGIT, OPROBIT und GOLOGIT, GOPROBIT für das MULTINOMIALE LOGIT Modell Beispiel Berufliche Stellung (Abschnitt 8.1)

```
-------------------------------------------------------------------------------
. name:  oprobit_ologit_brant_test
. log:  x:\EWF\_Discrete Choice\Beispiele\MNL estimates
. Stata\oprobit_ologit_brant_test.log
. log type:  text
. opened on:  29 Jun 2016, 12:19:12

. use MNLEXAMPLE-dat_V10, replace
.
. * install brant test und andere ordered Modelle (über GOOGLE downloadable)
. do brant_test_Pre_Dofile_Packages
```

```
…

.
. sum *

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
      observ |        30        15.5    8.803408          1         30
           y |        30         1.6    1.191927          0          3
      educat |        30    9.833333    1.913353          7         14
         sex |        30    .4666667    .5074163          0          1
      income |        30    2289.667    939.8926        750       4500

. global x educat sex

.
. * MNL
. mlogit y $x, base(0)

Iteration 0:   log likelihood = -40.942646
Iteration 1:   log likelihood =  -33.14412
Iteration 2:   log likelihood =  -32.64134
Iteration 3:   log likelihood = -32.628838
Iteration 4:   log likelihood = -32.628809
Iteration 5:   log likelihood = -32.628809

Multinomial logistic regression                 Number of obs   =         30
                                                LR chi2(6)      =      16.63
                                                Prob > chi2     =     0.0108
Log likelihood = -32.628809                     Pseudo R2       =     0.2031

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
0            |  (base outcome)
-------------+----------------------------------------------------------------
1            |
      educat |   .2136465   .5149648     0.41   0.678     -.795666    1.222959
         sex |   .9595302   1.180556     0.81   0.416    -1.354318    3.273378
       _cons |  -2.748282   4.510059    -0.61   0.542    -11.58783    6.091271
-------------+----------------------------------------------------------------
2            |
      educat |   1.248257   .5363583     2.33   0.020     .1970145      2.2995
         sex |   .3959491   1.236407     0.32   0.749    -2.027365    2.819263
       _cons |  -11.86602   4.967337    -2.39   0.017    -21.60182   -2.130219
-------------+----------------------------------------------------------------
3            |
      educat |   1.343701   .5423721     2.48   0.013     .2806712    2.406731
         sex |   .1559369   1.247926     0.12   0.901    -2.289953    2.601827
       _cons |  -12.67277   5.038821    -2.52   0.012    -22.54868   -2.796859
------------------------------------------------------------------------------

.
. *ordered probit
. oprobit y $x

Iteration 0:   log likelihood = -40.942646
Iteration 1:   log likelihood =  -34.78658
Iteration 2:   log likelihood = -34.757233
Iteration 3:   log likelihood = -34.757228
Iteration 4:   log likelihood = -34.757228

Ordered probit regression                       Number of obs   =         30
                                                LR chi2(2)      =      12.37
                                                Prob > chi2     =     0.0021
Log likelihood = -34.757228                     Pseudo R2       =     0.1511

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      educat |   .4232764   .1256828     3.37   0.001     .1769427    .6696101
         sex |   .0502119   .4135315     0.12   0.903    -.7602949    .8607187
-------------+----------------------------------------------------------------
```

```
       /cut1 |   3.331517   1.192241                      .994767    5.668266
       /cut2 |      3.964    1.24904                     1.515926    6.412074
       /cut3 |   4.876207   1.319411                      2.29021    7.462205
------------------------------------------------------------------------------

.
. * ordered logit
. ologit y $x

Iteration 0:   log likelihood = -40.942646
Iteration 1:   log likelihood = -34.632639
Iteration 2:   log likelihood = -34.478395
Iteration 3:   log likelihood = -34.477394
Iteration 4:   log likelihood = -34.477393

Ordered logistic regression                     Number of obs   =         30
                                                LR chi2(2)      =      12.93
                                                Prob > chi2     =     0.0016
Log likelihood = -34.477393                     Pseudo R2       =     0.1579

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      educat |    .741394   .2323991     3.19   0.001     .2859001    1.196888
         sex |   .2061211   .6927339     0.30   0.766    -1.151612    1.563855
-------------+----------------------------------------------------------------
       /cut1 |   5.975741   2.201242                      1.661385     10.2901
       /cut2 |   7.058016   2.327139                      2.496908    11.61912
       /cut3 |   8.662004   2.528519                      3.706199    13.61781
------------------------------------------------------------------------------

. brant

Brant Test of Parallel Regression Assumption

    Variable |      chi2   p>chi2    df
-------------+--------------------------
         All |      3.58    0.465     4
-------------+--------------------------
      educat |      2.97    0.226     2
         sex |      0.85    0.655     2
----------------------------------------

A significant test statistic provides evidence that the parallel
regression assumption has been violated.

.
. * generalized ordered logit
. gologit2 y $x

Generalized Ordered Logit Estimates             Number of obs   =         30
                                                LR chi2(6)      =      19.96
                                                Prob > chi2     =     0.0028
Log likelihood = -30.964807                     Pseudo R2       =     0.2437

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
0            |
      educat |   1.099031   .5033277     2.18   0.029     .1125265    2.085535
         sex |   1.035823   1.019082     1.02   0.309     -.961542    3.033187
       _cons |  -9.474623   4.445046    -2.13   0.033    -18.18675   -.7624927
-------------+----------------------------------------------------------------
1            |
      educat |   1.674204   .6794755     2.46   0.014     .3424567    3.005952
         sex |  -.5234157   1.139439    -0.46   0.646    -2.756676    1.709844
       _cons |  -15.33406    6.12612    -2.50   0.012    -27.34103   -3.327086
-------------+----------------------------------------------------------------
2            |
      educat |   .3360817   .2776213     1.21   0.226     -.208046    .8802095
         sex |   -.122005   .8850838    -0.14   0.890    -1.856737    1.612727
       _cons |  -3.971384   2.937733    -1.35   0.176    -9.729234    1.786467
------------------------------------------------------------------------------
```

```
WARNING! 15 in-sample cases have an outcome with a predicted probability that is
less than 0. See the gologit2 help section on Warning Messages for more information.

.
.
. * generalized ordered probit
. goprobit y $x

Iteration 0:   log likelihood = -40.942646
Iteration 1:   log likelihood = -32.379991
Iteration 2:   log likelihood = -31.023569
Iteration 3:   log likelihood = -30.874474
Iteration 4:   log likelihood =  -30.87264
Iteration 5:   log likelihood =  -30.87264

Generalized Ordered Probit Estimates              Number of obs   =         30
                                                  LR chi2(6)      =      20.14
Log likelihood =  -30.87264                       Prob > chi2     =     0.0026

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
mleq1        |
      educat |   .6362661      .2769     2.30   0.022      .093552     1.17898
         sex |   .5509966   .5919373     0.93   0.352    -.6091792    1.711172
       _cons |  -5.476655   2.451878    -2.23   0.026    -10.28225    -.671063
-------------+----------------------------------------------------------------
mleq2        |
      educat |   .9629437   .3591489     2.68   0.007     .2590247    1.666863
         sex |  -.3132307   .6655251    -0.47   0.638    -1.617636    .9911746
       _cons |  -8.833041   3.245935    -2.72   0.007    -15.19496   -2.471126
-------------+----------------------------------------------------------------
mleq3        |
      educat |   .2005889   .1610015     1.25   0.213    -.1149683    .5161462
         sex |  -.1014359    .539417    -0.19   0.851    -1.158674     .955802
       _cons |  -2.375314   1.673158    -1.42   0.156    -5.654643     .904014
------------------------------------------------------------------------------

.
end of do-file
```

11 Das POISSON-Modell für Zähldaten

11.1 Modellspezifikation und Schätzung

Wenn die abhängige Variable zwar diskrete Werte 0,1,2,3.... annimmt, dies aber keine kategorialen Werte sind, dann ist ein Zähldaten-Modell ('count model') für die Anzahl von Ereignissen angebracht.

POISSON-Modell

jede Beobachtung y_i ist die Realisation einer Poisson-Verteilung mit Parameter $\lambda = \lambda(x)$

$$P(Y_i = y_i) = \frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!} \qquad y_i = 0, 1, 2 \ldots$$

Für die Modellierung von λ_i wird meist das log-lineare Modell

$$\ln \lambda_i = x_i\beta \qquad (\lambda_i = e^{x_i\beta})$$

verwendet. Die erwartete Anzahl von Ereignissen ist

$$E[y_i \mid x_i] = Var[y_i \mid x_i] = \lambda_i = e^{x_i\beta} .$$

Marginale Effekte

$$\frac{\partial E[y_i \mid x_i]}{\partial x} = \lambda_i\beta$$

Schätzung über Maximum Likelihood (L)

$$\max_\beta(L) = \max_\beta(\ln L)$$

$$L = \prod_{i=1}^{n} P_i = \prod_{i=1}^{n} \frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!}$$

$$\ln L = \sum_{i=1}^{n} \ln P_i = \sum_i \left[-e^{x_i\beta} + y_i(x_i\beta) - \ln y!\right]$$

$$\frac{\partial \ln L}{\partial \beta} = \sum_i (-e^{x_i\beta}x_i + y_i x_i) = \sum_i (y_i - \lambda_i)x_i \overset{!}{=} 0$$

Newton-Raphson ergibt den Schätzer für β.

Die Hesse'sche Matrix

$$\frac{\partial^2 \ln L}{\partial \beta \, \partial \beta'} = \sum_i \lambda_i x_i x_i'$$

ist negativ definit (→ eindeutiges globales Optimum/Konvergenz). Asymptotische Kovarianzmatrix wieder über die Inverse der negativen Hesse'schen Matrix.

Modellgüte

z.B. über Likelihood Ratio Test, hier

$$LR = 2\sum_i \ln\left(\frac{\hat{P}_i(\beta)}{\hat{P}_i(0)}\right) \qquad \text{(LR ist } \chi^2 \text{ verteilt).}$$

Speziellere Maße: Greene 2008a, chapt. 25

11.2 Beispiele: POISSON-Modell

- Chemical entities (Holland 1973)
- Effects of R & D expenditures on patents (Hausman et al. 1981)
- Schiffsunfälle (Greene 1997, S. 932 f) (Tab. 6.2)

Tab. 6.2: Schiffsunfälle

The number of accidents per service month for a sample of ship types are listed in Table 19.14. The ships are of five types constructed in one of four periods. The observation is over two periods. Since ships constructed from 1975 to 1979 could not have operated from 1960 to 1974, there is one missing observation in each group. The second observation for group E is also missing, for reasons un-

TABLE 19.14 Ship Accident Data

Ship Type						*Year Constructed*				*Years Operated*		*Service*	*Accidents*	
Type	*A*	*B*	*C*	*D*	*E*	*60–64*	*65–69*	*70–74*	*75–79*	*60–74*	*75–79*	*Months*	*Actual*	*Fitted*
1	1	0	0	0	0	1	0	0	0	1	0	127	0	0.210
1	1	0	0	0	0	1	0	0	0	0	1	63	0	0.153
1	1	0	0	0	0	0	1	0	0	1	0	1095	3	3.638
1	1	0	0	0	0	0	1	0	0	0	1	1095	4	5.341
1	1	0	0	0	0	0	0	1	0	1	0	1512	6	5.673
1	1	0	0	0	0	0	0	1	0	0	1	3353	18	18.468
1	1	0	0	0	0	0	0	0	1	1	0	—	—	—
1	1	0	0	0	0	0	0	0	1	0	1	2244	11	8.516
2	0	1	0	0	0	1	0	0	0	1	0	44882	39	43.129
2	0	1	0	0	0	1	0	0	0	0	1	17176	29	24.229
2	0	1	0	0	0	0	1	0	0	1	0	28609	58	55.132
2	0	1	0	0	0	0	1	0	0	0	1	20370	53	57.624
2	0	1	0	0	0	0	0	1	0	1	0	7064	12	15.373
2	0	1	0	0	0	0	0	1	0	0	1	13099	44	41.847
2	0	1	0	0	0	0	0	0	1	1	0	—	—	—
2	0	1	0	0	0	0	0	0	1	0	1	7117	18	15.666
3	0	0	1	0	0	1	0	0	0	1	0	1179	1	0.981
3	0	0	1	0	0	1	0	0	0	0	1	552	1	0.674
3	0	0	1	0	0	0	1	0	0	1	0	781	0	1.303
3	0	0	1	0	0	0	1	0	0	0	1	676	1	1.656
3	0	0	1	0	0	0	0	1	0	1	0	783	6	1.475
3	0	0	1	0	0	0	0	1	0	0	1	1948	2	5.388
3	0	0	1	0	0	0	0	0	1	1	0	—	—	—
3	0	0	1	0	0	0	0	0	1	0	1	274	1	0.522
4	0	0	0	1	0	1	0	0	0	1	0	251	0	0.386
4	0	0	0	1	0	1	0	0	0	0	1	105	0	0.237
4	0	0	0	1	0	0	1	0	0	1	0	288	0	0.888
4	0	0	0	1	0	0	1	0	0	0	1	192	0	0.869
4	0	0	0	1	0	0	0	1	0	1	0	349	2	1.216
4	0	0	0	1	0	0	0	1	0	0	1	1208	11	6.177
4	0	0	0	1	0	0	0	0	1	1	0	—	—	—
4	0	0	0	1	0	0	0	0	1	0	1	2051	4	7.226
5	0	0	0	0	1	0	0	0	1	0	1	45	0	0.235
5	0	0	0	0	1	1	0	0	0	0	1	—	—	—
5	0	0	0	0	1	0	1	0	0	1	0	789	7	3.612
5	0	0	0	0	1	0	1	0	0	0	1	437	7	2.937
5	0	0	0	0	1	0	0	1	0	1	0	1157	5	5.982
5	0	0	0	0	1	0	0	1	0	0	1	2161	12	16.400
5	0	0	0	0	1	0	0	0	1	1	0	—	—	—
5	0	0	0	0	1	0	0	0	1	0	1	542	1	2.834

explained by the authors.[55] The substantive variables in the model are number of accidents in the observation period and aggregate number of service months for the ship type by construction year for the period of operation.

Estimates of the parameters of a Poisson regression model are shown in Table 19.15. The model is

$$\ln E[\text{accident per month}] = \boldsymbol{\beta}'\mathbf{x}.$$

The model therefore contains the ship type, construction period, and operation period effects, and the aggregate number of months with a coefficient of 1.0.[56] The model is shown in Table 19.15, with sets of estimates for the full model and with the model omitting the type and construction period effects. Predictions from the estimated full model are shown in the last column of Table 19.14.

The hypothesis that the year of construction is not a significant factor in explaining the number of accidents is strongly rejected by the likelihood ratio test.

$$\chi^2 = 2[84.11514 - 68.41455] = 31.40118.$$

TABLE 19.15 Estimated Poisson Regressions (Standard Errors in Parentheses)

Mean Dependent Variable 10.47

Variable	*Full Model*		*No Ship Type Effect*		*No Period Effect*	
Constant	−6.4029	(0.2175)	−6.9470	(0.1269)	−5.7999	(0.1784)
Type = A						
Type = B	−0.5447	(0.1776)			−0.7437	(0.1692)
Type = C	−0.6888	(0.3290)			−0.7549	(0.3276)
Type = D	−0.0743	(0.2906)			−0.1843	(0.2876)
Type = E	0.3205	(0.2358)			0.3842	(0.2348)
60–64						
65–69	0.6959	(0.1497)	0.7536	(0.1488)		
70–74	0.8175	(0.1698)	1.0503	(0.1576)		
75–79	0.4450	(0.2332)	0.6999	(0.2203)		
Period = 60–74						
Period = 75–79	0.3839	(0.1183)	0.3875	(0.1181)	0.5001	(0.1116)
Log service	1.0000		1.0000		1.0000	
Log L	−68.41455		−80.20123		−84.11514	
G^2	38.96262		62.53596		70.34967	
R_p^2	0.94560		0.89384		0.90001	
R_d^2	0.93661		0.89822		0.88556	

[55]Data are from McCullagh and Nelder (1983). See Example 8.2 for details.

[56]When the length of the period of observation varies by observation by T_i, and the model is of the rate of occurrence of events *per unit of time*, then the mean of the observed distribution is $T_i\lambda_i$. This produces the coefficient of 1.0 on the number of periods of service in the model.

The critical chi-squared value for three degrees of freedom is 7.82. The ship type effect is likewise significant,

$$\chi^2 = 2[80.20123 - 68.41455] = 23.57336,$$

against a critical value for four degrees of freedom of 9.49. The LM tests for the two restrictions give the same conclusions, but much less strongly. The value is 12.544 for the ship type effect and 8.750 for the period effects.

Quelle: Greene 1997, S. 933-935

12 Simultane diskrete Entscheidungsmodelle

12.1 Modellspezifikationen

Analog zu simultanen Systemen wäre eine Erweiterung der Eingleichungsmodelle hier um weitere Gleichungen mit korrelierten Störtermen (Greene 1997, chap. 19.6, Maddala 1983, chapt. 7, Amemiya 1985, S. 311 ff) wünschenswert.

Wir konzentrieren uns hier auf multivariate PROBIT-Modelle. Analog, aber wegen der geschlossenen Lösung einfacher, sind MULTIVARIATE LOGIT-Modelle zu bilden (vgl. Schmidt und Strauss 1975b).

Multivariat – im Gegensatz zu multinomial – heißt: es wird ein Mehrgleichungssystem von z.B. PROBIT-Gleichungen mit korrelierten Störtermen betrachtet.

Auf Gleichungssysteme mit interdependenten Gleichungen aus diskreten und kontinuierlichen 'linken Seiten' wird hier nicht eingegangen (vgl. dazu z.B. Pohlmeier 1989, Maddala 1983, chapt. 7).

12.2 BIVARIATES und MULTIVARIATES PROBIT

BIVARIATES PROBIT

Modell

$$\begin{aligned} y_1^* &= \beta_1^* x_1 + \varepsilon_1 \qquad y_1 = 1, \, wenn \; y_1^* > 0, \;\; 0 \; sonst \\ y_2^* &= \beta_2^* x_2 + \varepsilon_2 \qquad y_2 = 1, \, wenn \; y_2^* > 0, \;\; 0 \; sonst \\ E[\varepsilon_1] &= E[\varepsilon_2] = 0 \\ Var[\varepsilon_1] &= Var[\varepsilon_2] = 1 \\ Cov[\varepsilon_1 \varepsilon_2] &= \rho \end{aligned}$$

d.h. die Entscheidung für das Eine (eine Alternative) korreliert mit der Entscheidung für das Andere.

Schätzung

Bivariate Normalverteilung cdf

$$P(X_1 < x_1, X_2 < x_2) = \int_{-\infty}^{x_2}\int_{-\infty}^{x_1} \phi_2(z_1, z_2, \rho)\, dz_1 dz_2$$
$$= \phi_2(x_1, x_2, \varphi)$$
$$pdf \quad \phi_2(x_1, x_2, \rho) = \frac{\exp\left[-(1/2)\left(x_1^2 + x_2^2 - 2\rho x_1 x_2\right)/\left(1-\rho^2\right)\right]}{2\pi\left(1-\rho^2\right)^{1/2}}$$

Sei

$$q_{i1} = 2y_{i1} - 1 \quad und \quad q_{i2} = 2y_{i2} - 1, \; d.h.$$
$$q_{ij} = 1 \quad wenn \; y_{ij} = 1 \; und \; q_{ij} = -1 \quad wenn \quad y_{ij} = 0 \;\; (j = 1,2)$$
$$z_{ij} = \beta'_j x_{ij} \; und \; w_{ij} = q_{ij} z_{ij} \quad (j = 1,2)$$
$$\rho_{i*} = q_{i1} q_{i2} \rho$$

Wahrscheinlichkeit

$$P(Y_1 = y_{i1}, Y_2 = y_{i2}) = \phi_2(w_{i1}, w_{i2}, \rho_{i^{**}})$$

Log-Likelihood

$$\ln L = \sum_i \ln \phi_2(w_{i1}, w_{i2}, \rho_{i^*})$$

Bedingungen 1. Ordnung

$$\frac{\partial \ln L}{\partial \beta_j} = \sum_i \left(\frac{q_{ij} g_{ij}}{\phi_2}\right) x_{ij} = 0 \quad (j = 1,2)$$
$$\frac{\partial \ln L}{\partial \Phi} = \sum_i \left(\frac{q_{i1} q_{i2} \rho_2}{\phi_2}\right) = 0$$
$$wobei \quad g_{ij} = \Phi(w_{i1})\, \rho\left(\frac{w_{i2} - \rho_{i*} w_{i1}}{\sqrt{1-\rho_{i^*}^2}}\right)$$

ML-Lösung: β_1, β_2, ρ relativ aufwendig da in jeder Iteration für jede Betrachtung zweidimensionale Integrale zu berechnen sind.

Hesse'sche Matrix und damit Kovarianzmatrix (β_1, β_2, ρ) relativ komplex (vgl. Greene 1997, S. 908-909)

MULTIVARIATES PROBIT

Prinzipiell ist das multivariate Modell nur die Erweiterung des bivariaten Modells um weitere Gleichungen. Das praktische Problem dabei ist die Lösung mehrdimensionaler Integrale. Für

trivariate Integrale gibt es Ansätze, darüber hinaus sind diese Ansätze aber sehr aufwendig. Erst in neuerer Zeit sind dafür so genannte 'simulation estimators' entwickelt worden (McFadden 1999)

'Simulation estimators' Idee
Näherungsweise Lösung der mehrdimensionalen Integrale für die Berechnung einer Wahrscheinlichkeit (1. Ableitungen etc.) basiert auf einer Mittelung von R Ziehungen einer bestimmten multivariaten Normalverteilung für jede Beobachtung. Genauere Ergebnisse sind natürlich mit größerem R aber immer noch sehr computerzeitaufwendig erhältlich.

Hinweis: LIMDEP Version 7.0 arbeitet mit der GHK Simulationsmethode und erlaubt bis zu 20 (!) miteinander korrelierte PROBIT-Gleichungen (Greene 1998, S. 494)

12.3 Beispiel: BIVARIATES PROBIT - Voting Behaviour

Voting behaviour: BIVARIATE PROBIT

Greene (1984) reports estimates of a model of voter behavior in two decisions, whether to send at least one child to public school and whether to vote in favor of a school budget. Regressors in the model were'income', 'taxes' = property taxes paid, and 'years' of residence in the community. The results in Table 19.6 were obtained using 95 observations.

TABLE 19.6 Bivariate Probit Estimates

	Single Equation		*Bivariate*	
Public school				
Constant	−4.972	(3.996)	−4.764	(4.009)
Income	0.1537	(0.4495)	0.1149	(0.5958)
Taxes	0.6440	(0.5654)	0.6699	(0.7226)
School budget				
Constant	−0.6862	(4.009)	−0.3066	(3.850)
Income	0.9961	(0.4404)	0.9895	(0.4863)
Taxes	−1.2646	(0.5672)	−1.3080	(0.6748)
Years	−0.01644	(0.0147)	−0.0176	(0.0136)
ρ	0.0		0.137	(0.2303)
Log-likelihood	−40.0830 + −58.5006		−97.4117	

Quelle: Greene 1993, S. 660

12.4 Beispiel: MULTIVARIATES PROBIT - The Timing of Daily Demand for Goods and Services

Timing of Daily Demand for Goods and Services – Microsimulation Policy Results of an Aging Society, Increasing Labour Market Flexibility, and Extended Public Childcare in Germany (Merz, Hangelberger und Rucha 2010 and 2009 as FFB-Discussion Paper No. 77)

Goodness of Fit and the Importance of Simultaneous Over-the-day Demand

The timing of the daily demand (as a further used shortcut for the probability of the timing of daily demand for goods respective services) overall is significantly explained by our model (Likelihood Ratio test with α= 0.1%, see Table 3). The importance of simultaneity is visible by the final correlation matrix of the error terms of the Multivariate Probit estimation.[2] The respective Table 3 is organised in a goods' and services' block structure. Early and morning shopping as well as afternoon and evening shopping are highly correlated showing overlapping demand activities. This is true also for the demand for services. The interdependence between demanding goods and asking for services is strong for each daily period (highly significant positive correlation).

Two daily demand regimes, the whole morning till 1 p.m. and the whole afternoon till 8 p.m., will be visible with almost no overlapping demand activities. Thus shopping and the demand for services are done closely together either mornings or in the afternoon but not spread over the day. This strong picture for the active population (caused probably by working hour restrictions) is given to a certain extent for the non-active population, too. Though we face two prominent daily demand regimes we will see that the single different significances and differences of the demand period specification support the prominent importance of the chosen four daily demand period thresholds.

Table 3: Correlation Matrix for the Multivariate Probit estimation of the demand for goods and services by time of the day for the active population

	Goods 6-9	Goods 9-13	Goods 13-17	Goods 17-20	Services 6-9	Services 9-13	Services 13-17	Services 17-20
	ε_1	ε_2	ε_3	ε_4	ε_5	ε_6	ε_7	ε_8
ε_1	1	0.253 ***	-0.003	0.048	0.231***	0.114	-0.014	-0.012
ε_2		1	-0.008	-0.072 *	0.178***	0.293***	-0.072	-0.064
ε_3			1	0.285 ***	-0.161**	-0.261***	0.192***	-0.009
ε_4				1	0.013	-0.203***	0.026	0.096**
ε_5					1	0.386***	-0.021	0.063
ε_6						1	0.05	0.141*
ε_7							1	0.441***
ε_8								1
LR-Test : H_0: $V=I_8$			χ^2 (28)	518.4 ***				

* p-value<0,05 / ** p-value<0,01 / *** p-value<0,001; Source: German Time Budget Survey 2001/2002, own calculations

Die Schätzergebnisse (β) für die Wahrscheinlichkeit in einer Tageszeitperiode einzukaufen, bzw. Dienstleistungen nachzufragen finden sich in Merz, Hanglberger und Rucha 2010).

[2] Considering computing expenses: Our Simulated Maximum Likelihood estimates with eight dimensional integrals to be solved for each parameter and each iteration that time was quite time consuming with about 100 hours (full four days and nights) for a one group eleven iteration solution with respective 25 simulation draws and a PC with a Genuine INTEL 2160 1,8 GHz processor and 512 MB RAM.

IV MODELLE MIT BEGRENZT ABHÄNGIGEN VARIABLEN

1 Modellstruktur Limited dependent variable (LDV)-Modelle: truncation, censoring

Die zu erklärende Variable LHS (left hand side) ist nun nicht mehr diskret, sondern als beschränkt abhängige Variable (limited dependent variable) für bestimmte Bereiche kontinuierlich.

- Truncation: „abgeschnitten“
  es gibt für $y = f(x)$ weder für y noch für x Werte (Probanden) unterhalb/oberhalb eines bestimmten Schwellenwertes c;
  z.B. in eine Stichprobe gelangen nur Probanden mit einem Einkommen oberhalb einer Armutsschwelle c. Aus dieser Stichprobe wären dann die Einkommen für alle, Arme und Nicht-Arme, also für die gesamte Population zu schätzen.
- Censoring: „zensiert“
  für $y = f(x)$ gibt es zwar für alle Probanden x Werte aber y-Werte sind nur für bestimmte Bereiche vorhanden. Die y-Werte aus einem bestimmten Bereich sind transformiert zu (oder beobachtet als) einem bestimmten zensierten Wert c_y;
  z.B. *Ausgaben für dauerhafte Konsumgüter* (y): Es gibt Personen/Haushalte in einem Querschnitt, die z.B. ein Auto kaufen, andere wiederum nicht. Für die, die ein Auto kaufen gibt es eine große Spanne unterschiedlicher Ausgaben. Für die, die kein Auto kaufen, gibt es zwar keine entsprechenden Ausgaben ($y=c_y=0$), wohl aber Personen-/Haushaltsmerkmale mit entsprechenden x-Werten.
  z.B. *Arbeitsstunden* sind ein weiteres Beispiel und ein weiterer Anwendungsbereich zensierter Modelle (‚labour supply studies’): Auch die Nicht-Arbeitenden sind Bestandteil der Stichprobe (Arbeitszeit=c=0) und können durch x beschrieben und berücksichtigt werden.

Literatur: Greene 2000, chapter 19, Maddala 1983, chapter 6, Breen 1996).

Beispiele „zensierter“ (censored) und „abgeschnittener“ (truncated) Daten

Analysen mit zensierten Daten:

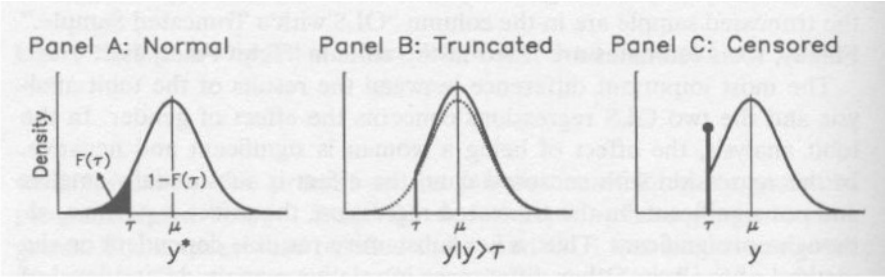
- Household purchases of durable goods (Tobin 1958)
- The number of extramarital affairs (Fair 1978)
- Female labour supply, Hours worked by a woman in the labor force (Quester und Greene 1982, Merz 1983)
- European mothers’ time spent looking after children (Joesch and Spiess 2006)

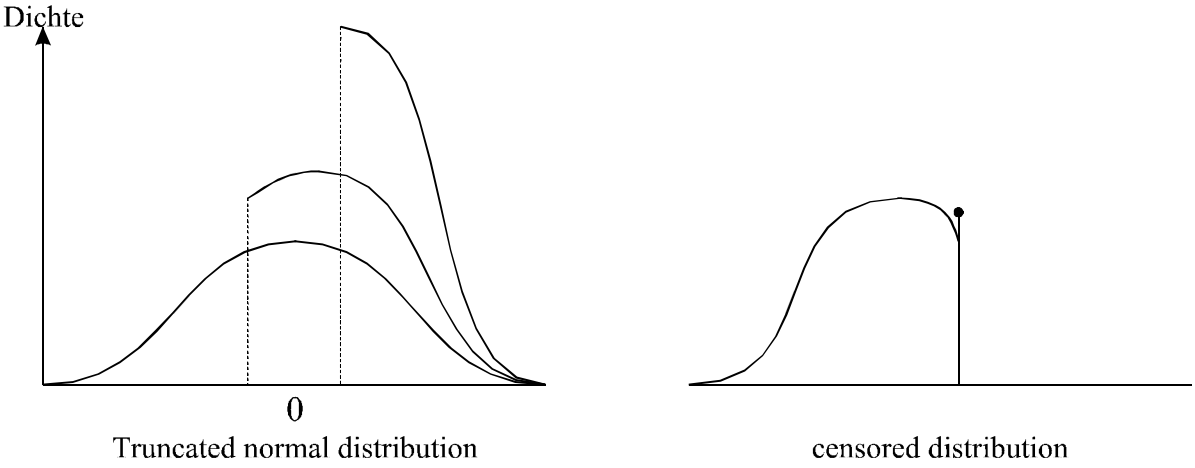
Allgemein „truncated“ und „censored“ Verteilung:

Truncated distribution: der abgeschnittene Teil der Dichtefunktion erhöht die Dichte im beobachteten Bereich

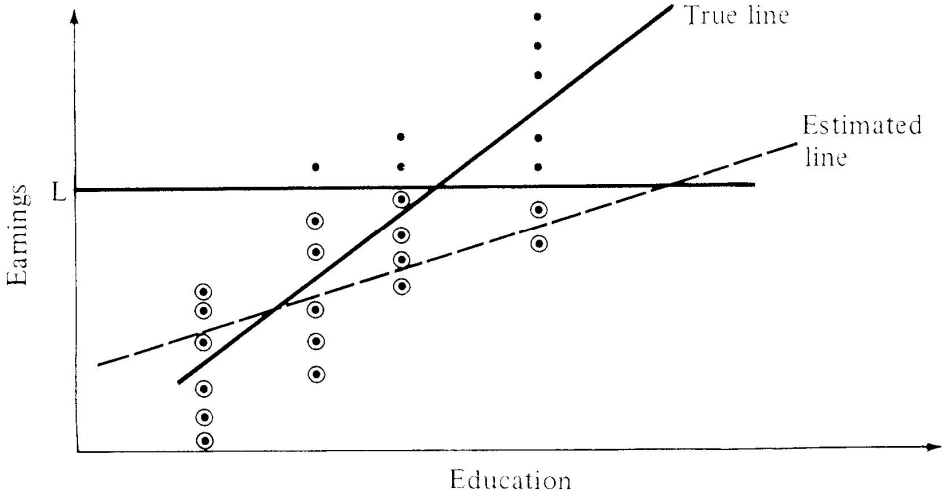
Censored distribution: der zensierte Teil der Verteilung fällt auf einen Punkt am Zensurwert c_y zusammen; die Wahrscheinlichkeitsdichten der anderen Werte bleiben unverändert.

Normal, Truncated and Censored Distributions (Long und Freese 1997, S. 192)



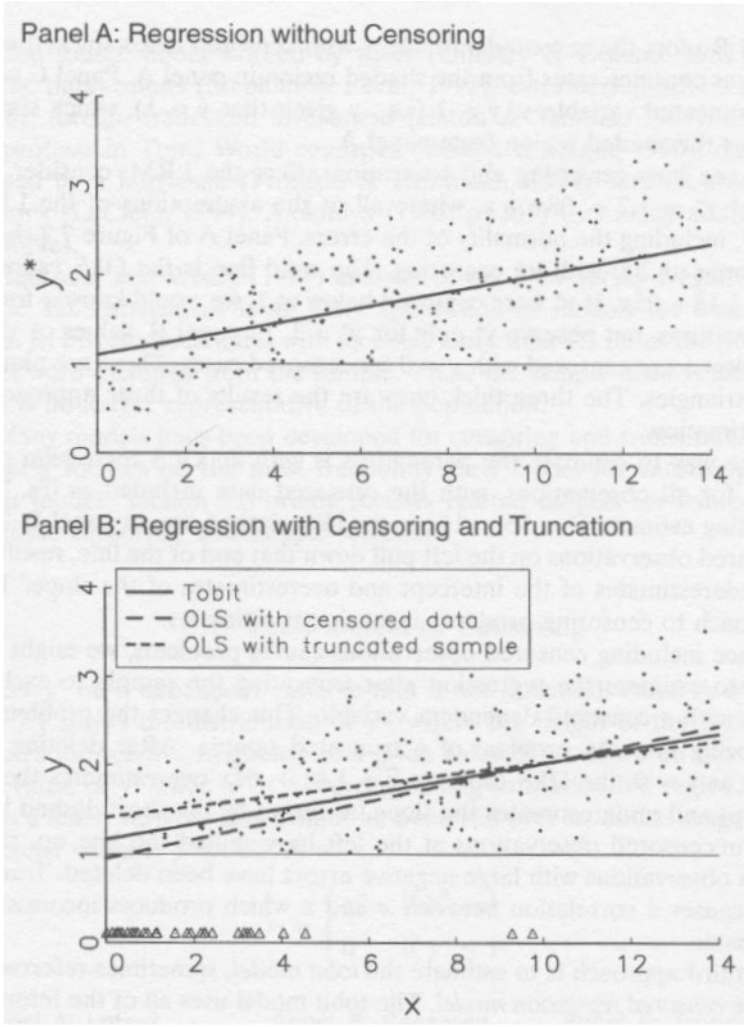


Schätzung aller Einkommen alleine aus den Daten armer Haushalte (truncated regression)



New Jersey Income Maintanance Experiment: all families with income > 1,5*official poverty line are truncated

Regression ohne Zensur, mit Zensur und Abschneiden (truncation)



(Long und Freese 1997, S. 190)

2 Das TOBIT-Modell

2.1 Modellspezifikation

Das TOBIT-Modell ist ein zensiertes Modell. Nach Greene 2003, S. 764 wird das TOBIT-Modell in Form eines Indexfunktionsmodells allgemein mit einer latenten Variablen y_i^* formuliert mit:

$$y_i^* = \beta' x_i + \varepsilon_i \qquad \text{'}underlying\ \ stochastic\ \ index\text{'}$$

$$y_i = \begin{cases} y_i^* & \text{wenn } y_i^* > c\ ,\ \ \varepsilon_i > -\beta' x_i \quad (oft\ c = 0) \\ 0 & \text{sonst } \left(y_i^* \le c\right) \end{cases}$$

Die Residuen ε_i sind unabhängig und normalverteilt mit $\left(E(\varepsilon_i) = 0, Var(\varepsilon_i) = \sigma^2\right)$.

Wird also durch die latente Bewertung (z.B. eine Art Kaufbereitschaft) ein Schwellenwert überschritten, dann ist die kontinuierliche Variable y beobachtbar. Die Lage des Schwellenwertes

(‚threshold') c ist aus dem Sachverhalt zu bestimmen. Häufig ergibt sich hierbei Null, es können sich aber auch andere Werte ergeben, die auch von Beobachtung zu Beobachtung verschieden sein können.

Es gibt also im TOBIT-Modell eine Anzahl von Beobachtungen mit zensiertem y-Wert, für die jedoch die erklärenden Größen x_i beobachtet wurden.

Der **Erwartungswert der latenten Variablen** y^* wegen $\varepsilon_i \sim N(0,\sigma)$ ist

$$Ey^* = \beta' x_i$$

Der **Erwartungswert aller y-Werte** unter Berücksichtigung aller x-Werte ist

$$Ey = \beta' x_i F(z_i) + \sigma f(z_i)$$

mit $z_i = \beta' x_i / \sigma$ und $f(z_i)$ als normalverteilte Dichtefunktion sowie $F(z_i)$ als die normalverteilte Verteilungsfunktion (cdf). Damit gehen sowohl die Partizipationswahrscheinlichkeit $F(z_i)$ als auch alle x-Werte in die erwarteten y- Werte ein.

Der **Erwartungswert der beobachteten Werte oberhalb des Schwellenwertes (** y^+) unter Berücksichtigung aller x-Werte ist

$$\begin{aligned} Ey_i^+ &= E(y_i \,|\, y_i > 0) \\ &= E(y_i \,|\, \varepsilon_i > -\beta' x_i) \\ &= \beta' x_i + \sigma f(z) / F(z) \end{aligned}$$

Neben $\beta' x_i$ ist also der Erwartungswert des abgeschnittenen (‚truncated') normalverteilten Fehlerterms ε_i mit $E\varepsilon_i = \sigma f(z) / F(z)$ zu berücksichtigen.

Damit ist eine grundlegende Beziehung zwischen dem **Erwartungswert aus allen y-Werten**, Ey_i , dem **Erwartungswert nur der y-Werte oberhalb des Schwellenwertes** (Ey_i^+) und der **Wahrscheinlichkeit oberhalb des Schwellenwertes zu sein** ($F(z_i)$) gegeben mit

$$Ey_i = F(z_i) Ey_i^+ .$$

Tobin's ursprüngliches Modell für Ausgaben dauerhafter Konsumgüter

Das TOBIT-Modell ist nach James Tobin (1958) – „Tobin's probit" nach Goldberger 1964 – benannt, der damit Konsumausgaben analysiert hat. Amemiya 1985 (S. 361 ff) bezeichnet obige Formulierung als das Standard Type I TOBIT-Modell.

Tobin (1958) hat sein Modell als ein **hybrides Modell** bezeichnet, das aus einem PROBIT-Modell und einem multiplen Regressionsmodell besteht. Er definiert sein Modell mit:

W „*limited dependent variable" (Ausgabenanteile)*

$Y_i = \beta' x_i$ "*linear combination of the independent variables to which W is by hypothesis related"*

Da Haushalte sich unterschiedlich verhalten, werden die weiteren Unterschiede durch eine (standardnormalverteilte) Zufallsvariable ε_i mit Mittelwert Null und Standardabweichung σ erfasst. Das Verhalten der Haushalte wird dann von Tobin beschrieben mit

$$W_i = \begin{cases} Y_i - \varepsilon_i & (Y_i - \varepsilon_i \geq L) \\ L & (Y_i - \varepsilon_i < L) \end{cases}.$$

Der Erwartungswert von W bei gegebenen Y und L wird abgeleitet als

$$E(W;Y,L) = LQ(.) + YP(.) + \sigma Z(.) \quad mit\,(.) = ((Y_i - L)/\sigma),$$

wobei die Partizipationswahrscheinlichkeit P eine normalverteilte Verteilungsfunktion (cdf), Q=1-P die Wahrscheinlichkeit unterhalb des Schwellenwertes zu liegen und Z eine normalverteilte Dichtefunktion ist. Damit gehen sowohl die Partizipationswahrscheinlichkeit als auch alle Werte in die erwarteten Ausgaben ein.

In Tobin's (1958) Beispiel werden die Ausgaben für dauerhafte Konsumgüter mit

$$W_i = \frac{\text{Ausgaben für dauerhafte Konsumgüter (durable goods expenditures)}}{\text{Verfügbares Einkommen (disposable income)}}$$

abhängig vom Alter (x_1=age of the head of the spending unit) und liquider Mittel (x_2=ratio of liquid asset holdings to total disposable income) untersucht.

Das ursprüngliche Modell von Tobin zielt auf die latente Variable $y_i^* = Y_i - \varepsilon_i$ mit den Beobachtungen $y_i = W_i$.

2.2 Interpretation des TOBIT-Modells – Zur Frage prinzipiell nicht negativer Beobachtungen (z.B. Ausgaben, Zeitverwendung etc.):

In vielen ökonometrischen Anwendungen wird alleine das Vorhandensein von Null-Beobachtungen schon als gegebene Voraussetzung für die Anwendung des (zensierten) Standard-TOBIT-Modells gesehen (es gibt nur nicht-negative beobachtete Werte und Werte mit Null): Positive Ausgaben für dauerhafte Konsumgüter für ein Teil der Stichprobe, für den anderen Teil der Stichprobe werden keine Ausgaben beobachtet (=Null); Arbeitsangebotsschätzungen mit positiven Arbeitszeiten oder nicht arbeitend (Arbeitszeit=Null); Zeitverwendung für Kinderbetreuung mit positiven Kinderbetreuungszeiten oder nicht betreuend (Kinderbetreuungszeit=Null).

Aber: Das Standard TOBIT-Modell nimmt an, dass die abhängige Variable y_i^* bei Null zensiert ist, es aber inhaltlich zu interpretierende y_i^* –Variablen (Anspruchsniveaus, Präferenzen, Bereitschaft zu …) auch im zensierten negativen Bereich gibt.

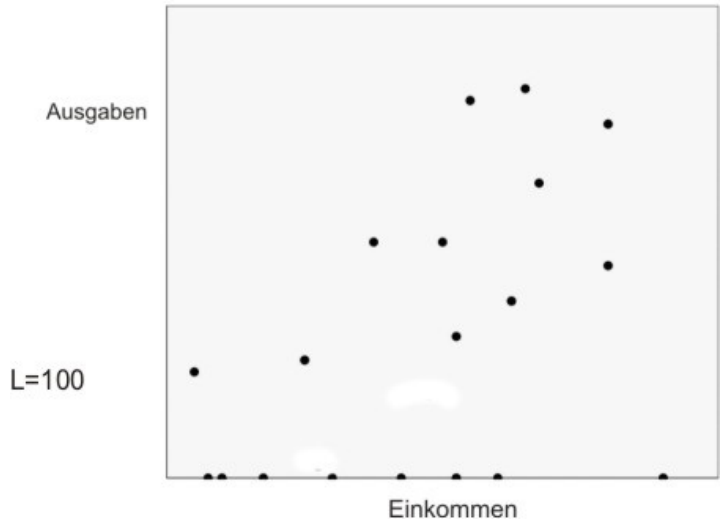
„ Every time we have zero observations in the sample, it is tempting to use the Tobit model. However, it is to understand what the model … really says. What we have … is a situation where y_i^* can, in principle, take on negative values. However, we do not observe them because of censoring. Thus the zero values are due to nonobservability. This is not the case with automobile expenditures, hours worked, or wages. These variables cannot, in principle, assume negative values. The observed zero values are due not to censoring, but due to the decisions of individuals. In this case the appropriate procedure would be to model the decisions that produce the zero observations rather than use the TOBIT model mechanically."

(Maddala 1992, S. 335).

Illustration 1: Ausgaben für dauerhafte Konsumgüter

Das klassische Tobin Beispiel: Ein Haushalt maximiere seinen Nutzen U(y,z) aus Ausgaben für dauerhafte Konsumgüter y (z.B. Automobil) und Ausgaben für alle anderen Konsumgüter z unter der Budgetrestriktion y+z<=Einkommen und der Begrenzung y>= Preis des am günstigsten zur Verfügung stehenden dauerhaften Konsumgutes. y*, die Bereitschaft bei Gelegenheit Geld auszugeben, sei die Lösung des Nutzenmaximierungsprozesses alleine aus der Budgetrestriktion.

TOBIT-Beispiel: Zensierte Ausgaben



Sei $y_i^* = \beta_0 + \beta_1 Einkommen + \varepsilon_i$ mit ε_i als Fehlerterm, der alle unbeobachtbaren Einflüsse auf die Nutzenfunktion beinhalte. Dann kann die Lösung des Ausgangsproblems, mit y als die Ausgaben für dauerhafte Konsumgüter, geschrieben werden als

$$y_i^* = \beta_0 + \beta_1 Einkommen + \varepsilon_i$$

$$y_i = \begin{cases} y_i^* = \beta_0 + \beta_1 Einkommen + \varepsilon_i & \text{wenn } y_i^* > L \text{ , } \varepsilon_i > -\beta' x_i \\ 0 & \text{sonst } \left(y_i^* \le L \right) \end{cases}$$

mit L=100, für alle billigstes Konsumgut (Auto)

Hat ein Konsument nur 50 € zur Verfügung und das günstigste dauerhafte Konsumgut (Auto) kostet 100 €, dann kann der Konsument dieses dauerhafte Konsumgut nicht kaufen. Er kann kaufen, wenn er mehr als 100 € hat. Beobachtet wird die Ausgabe, also der bezahlbare Preis des am günstigsten zur Verfügung stehenden Konsumgutes. Ist der Preis nicht bezahlbar, dann gibt es keine Ausgabegelegenheit, diese Fälle sind die zensierten Fälle. Die Ausgaben für dauerhafte Konsumgüter sind somit zensiert, da wir nicht wissen, wie viel der Haushalt ausgegeben hätte, wenn es ein dauerhaftes Konsumgut für weniger als 100 € geben würde.

Zusammengefasst: Die Ausgabebereitschaft wird durch das nicht vorhandene Angebot beschränkt, daraus folgt Zensur.

Illustration 2: Ticket Nachfrage und Ticket-Verkauf

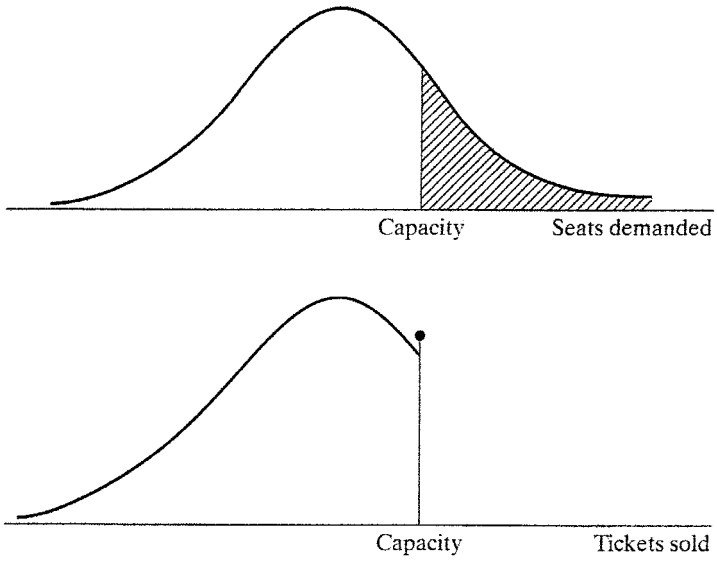
Beobachtet (gemessen) ist die Anzahl der aktuell verkauften Tickets. Eine Nachfrage bestehe auch dann noch, wenn ein Stadion (bspw.) bereits ausverkauft ist. Die Nachfrage nach Tickets ist zensiert und transformiert auf die Kapazitätsgrenze der verkauften Tickets (Zensierung „von oben"); wir wissen aber, dass die Nachfrage größer ist.

Die Interpretation der zensierten Schätzergebnisse (TOBIT) kann zweierlei sein:

Soll nun für ein kommendes Event die zu verkaufenden Tickets vorhergesagt werden, dann ist der zensierte Mittelwert (Ey) von Interesse.

Soll der Bedarf für den Bau eines neuen Stadions ermittelt werden, dann ist der Mittelwert der latenten Variablen (Ey^*) von Interesse.

FIGURE 22.2 Partially Censored Distribution.



Quelle: Greene 2003, S. 762

Illustration 3: Analyse der Arbeitszeit

Es gibt nur Jobs ab c Wochenstunden, die Anzahl der gewünschten Wochenstunden ist die latente Variable. Ist die gewünschte kleiner als die angebotene Arbeitszeit, dann wird keine Arbeit angenommen. Zensur der gewünschten Arbeitszeit erfolgt durch die vorhandene Arbeitskraftnachfrage, beobachtet werden nur zustande gekommene Arbeitsverhältnisse mit positiven Arbeitszeiten (y-Werte); TOBIT wäre angebracht.

Illustration 4: Zeitverwendung

In Zeittagebüchern (time use diaries) – wie aus den beiden deutschen Zeitbudgetstudien 1991/92 und 2001/02 – geben die Befragten für drei Tage in Intervallen von 10 Minuten an, welche Aktivität sie in dieser Zeit getätigt haben. Will man nun die verwendete Zeit über den gesamten Tag z.B. für die Analyse der Kinderbetreuung, Hausarbeit, tägliche Arbeitszeit – im Grunde für jede Aktivität – regressionsanalytisch erklären, so ist die nicht negative Ausprägung von Zeit, bzw. ein Wert von Null zu berücksichtigen. Man könnte von zensierten Werten sprechen und den TOBIT-Ansatz wählen, wenn ein kurzer Zeitraum (Zeit-Tagebuchtage) betrachtet wird und darin die zu erklärende Variable den zensierten Wert (z.B. keine Arbeitszeitangabe an einem der 3 Zeittagebücher einer Umfrage) annimmt. Bedenkt man aber, dass die jeweilige Aktivität über einen längeren Zeitraum eigentlich immer ausgeübt wird (z.B. es werden nur Arbeitende oder Eltern betrachtet), dann wäre OLS nur der nicht zensierten Werte angebracht (TOBIT würde die Koeffizienten „nach unten" verzerren (Stewart 2009).

Zur Frage „einfaches OLS vs. TOBIT bzw. zweistufige Schätzung"

Effizient ist es generell, alle vorhandenen Informationen auszunutzen, also sowohl die beobachteten erklärenden Werte oberhalb des Schwellenwertes, als auch die Information über die Nicht-Käufer, Nicht-Arbeitenden etc. Ein einfacher OLS-Ansatz, der nur die x-Werte mit positiven y-Werten verwendet verzichtet auf eine bessere Schätzung und Inferenz.

Als Alternative zum TOBIT-Ansatz kann der noch zu behandelnde zweistufige Schätzansatz von Heckman („two-part model") herangezogen werden, der in einer ersten Stufe aus allen Beobachtungen die Wahrscheinlichkeit schätzt, positive y-Werte zu haben. Diese Information geht über eine Selektionskorrektur (Mills-Ratio) dann zusätzlich in die OLS-Schätzung mit nur x-Werten mit gleichzeitigen positiven y-Werten ein und sichert unverzerrte Ergebnisse. Vorteil: Partizipation und Erklärung des kontinuierlichen Teils (z.B. Arbeitszeit) können getrennt erklärt werden. Ob eine Selektionsverzerrung vorliegt, also keine zufällige Auswahl der positiven y-Werte, kann im nachhinein dann über den geschätzten Koeffizienten der Selektionskorrektur gesehen werden.

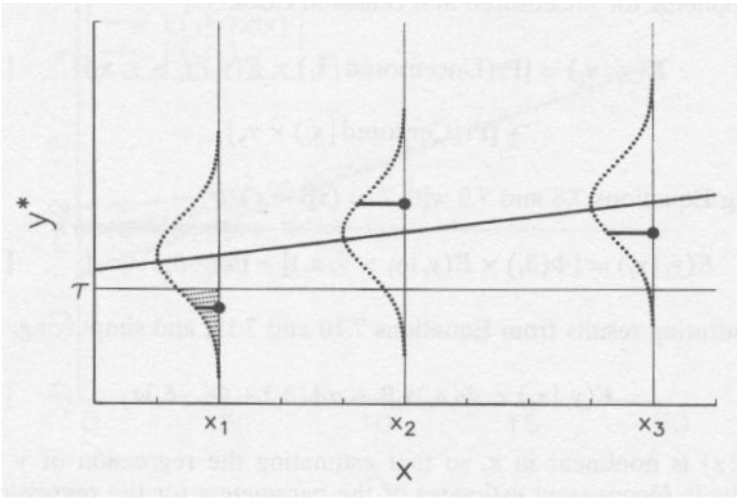
Daraus folgt, selbst wenn prinzipiell keine negativen (bspw.) Ausgaben vorkommen bzw. Zeit für eine Aktivität über einen längeren Zeitraum immer eingesetzt wird, so wäre OLS angebracht; und: es kommt wie immer auf die Interpretation an.

2.3 ML-Schätzung

Für die Maximum Likelihood Schätzung können die Beobachtungen in zwei Gruppen eingeteilt werden. Die in der ersten Gruppe enthaltenen Fälle – unzensierte Beobachtungen mit $y_i > 0$ –

werden, wie in der ML-Schätzung im linearen Regressionsmodell, die jeweiligen Dichtefunktionswerte in die Likelihood Funktion eingesetzt (entspricht der Höhe über den runden Punkten bei x_2 und x_3). Bei der zweiten Gruppe, den zensierten Fällen mit $y = 0$, liegt kein konkreter Wert für y_i^* vor. Stattdessen kann nur die gesamte Wahrscheinlichkeit überhaupt zensiert zu sein, berechnet und in die Likelihood Funktion eingesetzt werden (entspricht dem Fall x_1 in der nachfolgenden Grafik; die Wahrscheinlichkeit überhaupt zensiert zu sein, entspricht der gesamten Fläche unter der Dichtefunktion bis zum Schwellenwert τ) (Long und Freese 1997, S. 204).

Maximum Likelihood Schätzung für das TOBIT-Modell



Normalverteilung

Verteilungsfunktion cdf

$$F_i = F\left(\beta' x_i, \sigma^2\right) = \int_{-\infty}^{\beta' x_i} \frac{1}{\sigma (2\pi)^{1/2}} \exp\left(\frac{-t^2}{2\sigma^2}\right) dt$$

Dichtefunktion pdf

$$f_i = f\left(\beta' x_i, \sigma^2\right) = \frac{1}{\sigma (2\pi)^{1/2}} \exp\left[-\frac{(\beta' x_i)^2}{2\sigma^2}\right]$$

Standardnormalverteilung

Verteilungsfunktion cdf

$$\Phi_i = F_i = \int_{-\infty}^{\frac{-\beta' x_i}{\sigma}} \frac{1}{(2\pi)^{1/2}} \exp\left(\frac{-t^2}{2}\right) dt$$

Dichtefunktion pdf

$$\phi_i = \sigma f_i = \frac{1}{(2\pi)^{1/2}} \exp\left[-(\beta' x_i)/2\sigma^2\right]$$

$$\left(y = \beta' x_i / \sigma \sim N(0,1)\right)$$

Wahrscheinlichkeit für eine zensierte Beobachtung: $\text{Prob}\left(\text{y}_\text{i}=0\right)$

für $y_i=0$: $\text{Prob}\left(\text{y}_\text{i}=0\right)=\text{Prob}\ (\text{y}_\text{i}^*<0)=\text{Prob}\left(\varepsilon_i<-\beta' x_i\right)=1-F_i$

Prob(zensierte Beobachtung)=Prob(y_i^*<c)=$\Phi[(c-\beta' x_i)/\sigma]$

Wahrscheinlichkeit für eine nicht zensierte Beobachtung: $\text{Prob}\left(\text{y}_\text{i}>0\right)$

$$\text{für } y_i>0:\ \text{Prob}\left(y_i>0\right)\cdot f\left(y_i\middle|y_i>0\right)=F_i\frac{f\left(y_i-\beta' x_i,\sigma^2\right)}{F_i}$$

$$=\frac{1}{\left(2\pi\sigma^2\right)^{1/2}}\exp\left[-\left(\frac{1}{2\sigma^2}\right)\left(y-\beta' x_i\right)^2\right]$$

Prob(nicht zensierte Beobachtung)=1- $\Phi[(c-\beta' x_i)/\sigma]=\Phi[(\beta' x_i-c)/\sigma]$

TOBIT-Likelihood Funktion

$$L=\Pi_0\left(1-F_i\right)\Pi_1 f_i\left[\left(y_i-\beta' x_i\right)/\sigma\right]$$
$$=\Pi_0 F_i\left(y_{0i}\right)\Pi_1 f_i\left(y_i\right)$$
$$=\Pi_0\left[1-\Phi\left(\beta' x_i/\sigma\right)\right]\ \ \Pi_1\sigma^{-1}\phi\left[\left(y_i-\beta' x_i\right)/\sigma\right]\quad (censored)$$

$$[L=\Pi_1\Phi\left(\beta' x_i/\sigma\right)^{-1}\sigma^{-1}\phi\left[\left(y_i-\beta' x_i\right)/\sigma\right]\quad (truncated)]$$

L = Probit likelihood * truncated Tobit likelihood:

$$L=\Pi_0\left[1-\Phi\left(\beta' x_i/\sigma\right)\right]\Pi_1\Phi\left(\beta' x_i/\sigma\right)\ \ \Pi_1\Phi\left(\beta' x_i/\sigma\right)^{-1}\sigma^{-1}\phi\left[\left(y_i-\beta' x_i\right)/\sigma\right]$$

$$\log L=\sum\nolimits_0\log\left(1-F_i\right)+\sum\nolimits_1\log\left(\frac{1}{\left(2\pi\sigma^2\right)^{1/2}}\right)-\sum\nolimits_1\frac{1}{2\sigma^2}\left(y_i-\beta' x_i\right)^2$$

1. Ableitungen

$$\frac{\partial\log L}{\partial\beta}=-\sum\nolimits_0\frac{f_i x_i}{1-F_i}+\frac{1}{\sigma^2}\sum\nolimits_1\left(y_i-\beta' x_i\right)x_i=0$$

$$\frac{\partial\log L}{\partial\sigma^2}=\frac{1}{2\sigma^2}\sum\nolimits_0\frac{\beta' x_i f_i}{1-F_i}-\frac{n_1}{2\sigma^2}+\frac{1}{2\sigma^4}\sum\nolimits_1\left(y_i-\beta' x_i\right)^2=0\quad \text{(Maddala 1991, S. 153)}$$

Lösung nach Umformung

$$\beta = \beta_{OLS} - \sigma \left(X_1 X_1 \right)^{-1} X_0' y_0$$

wobei β_{OLS} der OLS-Schätzer allein aus den n_1 („Nicht-Null“, 1) Beobachtungen ist.

Varianz-Kovarianz Matrix V von β

Die negative Inverse der Informations-Matrix der 2. Ableitungen an der Stelle der (optimalen) Parameterschätzung ist wieder die asymptotische Varianz-Kovarianzmatrix. Damit kann dann die Signifikanz der Parameter β mit der Standardabweichung als Wurzel der entsprechenden Elemente der Diagonalen von V errechnet werden..

Direkte Berechnung der asymptotischen Varianz-Kovarianz Matrix V (Amemiya 1985, S. 373).

$$V\left(\Theta = \left(\beta', \sigma^2\right)\right) = \begin{bmatrix} \sum_i a_i x_i x_i' & \sum_i b_i x_i \\ \sum_i b_i x_i' & \sum_i c_i \end{bmatrix}^{-1}$$

wobei

$$a_i = -\sigma^2 \left\{ x_i' \alpha \phi_i - \left[\phi_i^2 / (1 - \phi_i) \right] - \Phi_i \right\}$$

$$b_i = (1/2)\sigma^{-3} \left\{ \left(x_i' \alpha\right)^2 \phi_i + \phi_i - \left[(x_i' \alpha) \phi_i^2 / (1 - \Phi_i) \right] \right\}$$

$$c_i = -(1/4)\sigma^{-4} \left\{ \left(x_i' \alpha\right)^3 \phi_i + \left(x_i' \alpha\right) \phi_i - \left[(x_i' \alpha) \phi_i^2 / (1 - \Phi_i) \right] - 2\Phi_i \right\}$$

ϕ_i, Φ_i ist $\phi(x_i' \alpha)$ und $\Phi(x_i' \alpha)$

$$\alpha = \beta / \sigma$$

Der nichtlineare TOBIT-Schätzer muss iterativ gelöst werden. Olsen (1978) hat bewiesen, dass log *L* global konkav ist, so dass Standard Newton-Raphson bspw. auch zu einem eindeutigen Maximum führt.

TOBIT-Eigenschaften

- TOBIT ML-Schätzung bleibt konsistent bei serieller Korrelation, aber nicht mehr konsistent bei Heteroskedastizität und Nicht-Normalität der Residuen
- Classical linear regression CLR: OLS (ML-Schätzung) bei Normalität bleibt generell konsistent bei den drei genannten ‚nonstandard‘ Annahmen.

Marginale Effekte

$$\frac{\partial E\left(y_i \middle| x_i\right)}{\partial x_i} = \beta \Phi\left(\frac{\beta' x_i}{\sigma} \right)$$

Dekomposition der Effekte bei einer marginalen Änderung von x_i

McDonald und Moffitt (1980) haben gezeigt, dass sich der Effekt einer marginalen Änderung einer erklärenden Variablen wie folgt aufteilen lässt:

$$\partial Ey_i / \partial x_i = F(z_i)(\partial Ey_i^+ / \partial x_i) + Ey_i^+(\partial F(z_i)/\partial x_i) \ .$$

Der gesamte marginale Effekt setzt sich damit aus zwei intuitiven Teilen zusammen: aus

1. *der Änderung der y-Werte oberhalb des Schwellenwertes* $\partial Ey_i^+ / \partial x_i$ gewichtet mit der Wahrscheinlichkeit oberhalb des Schwellenwertes zu sein $F(z_i)$ und

2. *der Änderung der Wahrscheinlichkeit oberhalb des Schwellenwertes* zu sein $\partial F(z_i)/\partial x_i$ gewichtet mit dem Erwartungswert der y-Werte oberhalb des Schwellenwertes Ey_i^+ .

$$\frac{\partial E(y_i|x_i)}{\partial x_i} = \beta[\Phi(1-\lambda_i(\alpha_i+\lambda_i))+\phi_i(\alpha_i+\lambda_i)]$$

wobei $\Phi_i = \Phi(\beta' x_i/\sigma) = \Phi(\alpha)$, $\lambda_i = \phi_i/\Phi_i$

Der Vektor der 1. Ableitungen zerlegt sich also in (Greene 1991, S.695)

$$\frac{\partial E(y_i|x_i)}{\partial x_i} = \text{Prob}\left(y_i^* > 0\right)\frac{\partial E\left(y_i^*|x_i, y_i^* > 0\right)}{\partial x_i} + E\left(y_i^*|x_i, y_i^* > 0\right)\frac{\partial \text{Prob}\left(y_i^* > 0\right)}{\partial x_i}$$

und damit in zwei Effekte: auf das bedingte Mittel von y_i^* im positiven Teil der Verteilung sowie auf die Wahrscheinlichkeit, dass die Beobachtung in diesen positiven Teil fällt.

Exkurs: OLS-Schätzung zensierter Daten (TOBIT-Modell)

Schätzung mit nur positiven y-Werten (vgl. Amemiya 1985, S. 366-368)

$$E\left(y_i \middle| y_i > 0\right) = \beta' x_i + \sigma\lambda(\beta' x_i/\sigma) \qquad \lambda(z) = \phi(z)/\Phi(z)$$

=> OLS ist verzerrt und inkonsistent

Schätzung mit Nullen und positiven y-Werten

$$E\left(y_i\right) = \Phi(\beta' x_i/\sigma)\beta' x_i + \sigma\phi(\beta' x_i/\sigma)$$

=> OLS ohne den zweiten Term wäre verzerrt; mehr dazu: siehe zweistufigen Heckman-Schätzer

Exkurs Ende.

2.4 Gütemaße und Spezifikationstests

Neben dem Likelihood-Ratio Test können Gütemaße analog zum Bestimmtheitsmaß berechnet werden, aber mit

$$E\left(y_i \middle| y_i^* > 0\right) = \beta' x_i + \sigma\lambda(-\beta' x_i / \sigma)$$

mit $\lambda(z) = \dfrac{\phi(z)}{1-\Phi(z)}$

Ronning 1991, S. 138, Veall und Zimmermann 1990

Pseudo R^2 (In Anlehnung an McKelvey und Zavoina's PROBIT R^2):

$$R^2_{MZ} = \frac{\sum_i \left(\hat{y}_i^* - \bar{\hat{y}}^*\right)^2}{\sum_i \left(\hat{y}_i^* - \bar{\hat{y}}^*\right)^2 + n\hat{\sigma}^2}$$

Dieses Maß hat sich im Vergleich mit 9 anderen Maßen (darunter Mc Faddens Pseudo R^2) als relativ bestes Gütemaß für das TOBIT-Modell herausgestellt ((Veall und Zimmermann 1994).

Als Spezifikationstest wird in der Literatur vor allem ein Test auf Heteroskedastizität betrachtet (Greene 1991, S. 698 ff, Maddala 1991, S. 178 ff, Ronning 1991, S. 140 ff).

2.5 Beispiele

- Applications of the standard TOBIT-Model (Tab. 2.4a)
- Konsumausgaben, Merz 1983 (Tab. 2.4b)
- Female Labor Supply, Merz 1990b (Tab. 2.4c)
- Greene LIMDEP: Women's labor supply (Tab. 2.4d)

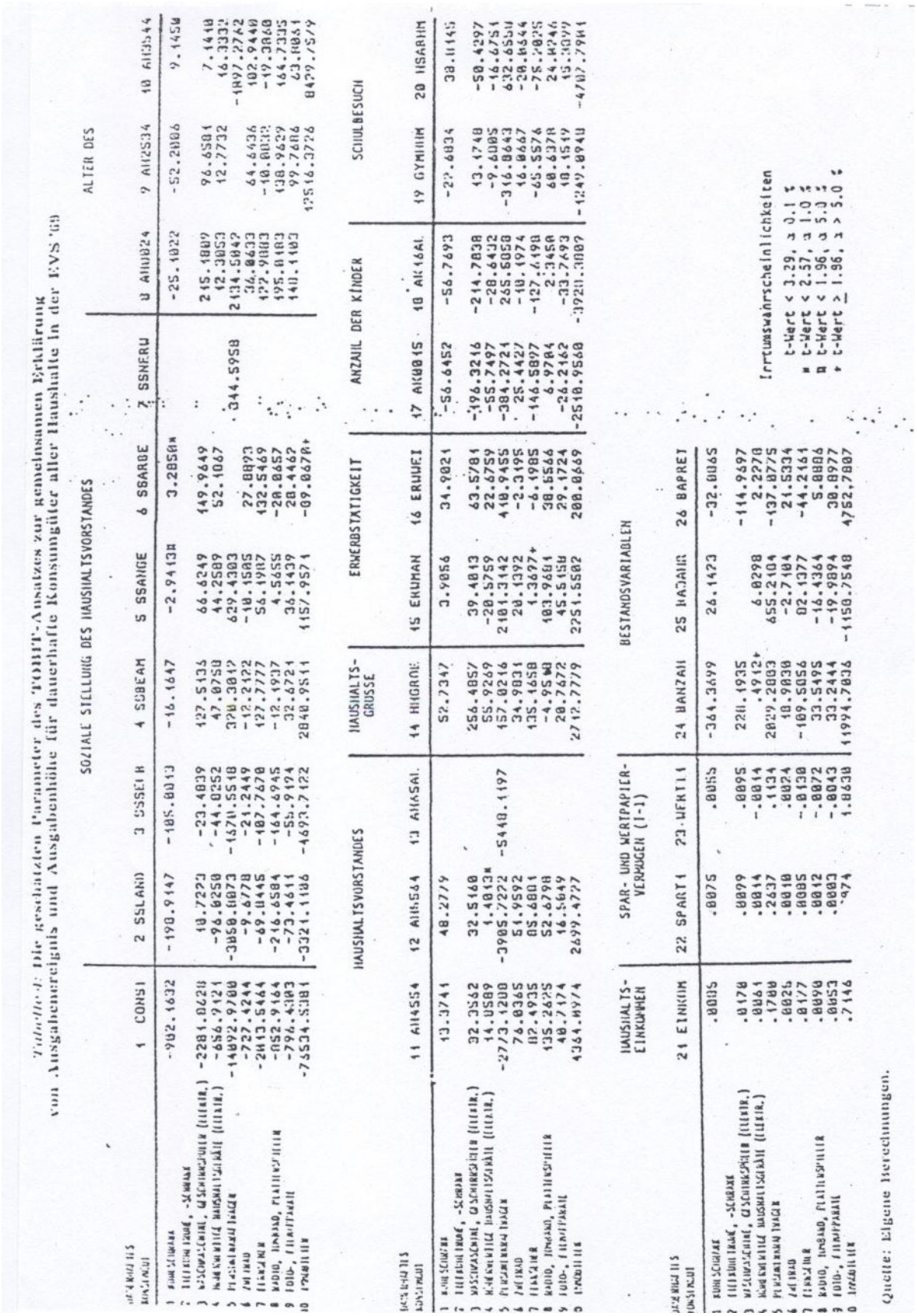
Tab. 2.4a: Early Applications of the Standard TOBIT-Model

Reference	y	x
Adams (1980)	Inheritance	Income, marital status, number of children
Ashenfelter and Ham (1979)	Ratio of unemployed hours to employed hours	Years of schooling, working experience
Fair (1978)	Number of extramarital affairs	Sex, age, number of years married, number of children, education, occupation, degree of religiousness
Keeley et al. (1978)	Hours worked after a negative income tax program	Preprogam hours worked, change in the wage rate, family characteristics
Kotlikoff (1979)	Expected age of retirement	Ratio of social security benefits lost at time of full-time employment to full.time earnings
Merz (1983)	Durable goods expenditures	Household head's characteristics, regional indicativs
Merz (2012, 1989)	Labor supply (hours of work) in market and non-market activities	Personal, household and regional characteristics
Merz (1990)	Female labor supply (working hours)	Personal, household and regional characteristics including partner's situation

Reece (1979)	Charitable contributions	Price of contributions, income
Rosenzweig (1980)	Annual days worked	Wages of husbands and wives, education of husbands and wives, income
Stephenson and McDonald (1979)	Family earnings after a negative income tax program	Earnings before the program, husband's and wife's educations, other family characteristics, unemployment rate, seasonal dummies
Wiggins (1981)	Annual marketing of new chemical entities	Research expenditure of the pharmaceutical industry, stringency of government regulatory standards
Witte (1980)	Number of arrests (or convictions) per month after release from prison	Accumulated work release funds, number of months after release until first job, wage rate after release, age race, drug use, drug use

Quelle: Amemiya 1983, S. 365

Tab. 2.4b: Konsumausgaben



Quelle: Merz 1983, S. 247, siehe auch Abschnitt II.6

Tab. 2.4c: Female Labor Supply

Table 6: Weekly working hours and labor force participation of married and unmarried women in the FRG.

	participation, working hours (TOBIT)			
	married		unmaried	
	coefficient	t-value[1]	coefficient	t-value
PERSONAL CHARACTERISTICS				
'Imputed' wage rate (ln hourly wage rate)[2]	61.122	(38.32)**	39.053	(21.82)**
Family status				
separate (married)			0.469	(0.10)
divorced			1.863	(0.74)
widowed			0.540	(0.44)
Age				
Age	-0.473	(1.31)	-0.189	(0.55)
Age^2*10^{-2}	0.088	(0.21)	-0.489	(1.28)
Health situation				
Impairment	3.224	(3.07)**	1.874	(1.12)
Chronic complaints[2]	-0.966	(0.63)	-2.671	(1.17)
(Subjective) satisfaction[3]				
Household income	-0.135	(0.55)	0.472	(1.54)
Household activities	-0.234	(0.91)	-0.012	(0.03)
Monthly unearned income, transfer payments				
Old age/widow's pension	-0.029	(3.47)**	-0.016	(5.10)**
Private support	-0.008	(0.35)	-0.021	(3.59)**
Unemployment benefit/relief (t-1)	-0.002	(0.50)	-0.015	(2.95)**
HUSBAND'S CHARACTERISTICS				
Occupational status				
Manual worker	1.851	(0.87)		
White-collar worker	-11.021	(4.53)**		
Civil-servant	-20.070	(6.60)**		
Self-employed	9.154	(3.48)**		
Net earned income last month	0.002	(4.43)**		
HOUSEHOLD CHARACTERISTICS				
Nursing cases existing	2.197	(0.58)	7.105	(1.70) *
Children in age classes				
0 to 5 years	-8.228	(6.61)**	-5.024	(1.44)
6 to 9 years	-2.524	(1.91) *	-0.078	(0.03)
10 to 15 years	-1.098	(1.22)	2.688	(1.78) *
16 years and older	0.374	(0.50)	1.903	(2.41)**
Monthly unearned income, transfer payments				
Household benfits	-0.004	(0.17)	-0.053	(2.00) *
Welfare benfits	-0.008	(0.51)	-0.016	(1.46)
Yearly unearned income: property				
Income from let and lease$*10^{-3}$	-0.747	(4.30)**	-0.473	(2.53)**
Income from interest and dividends$*10^{-3}$	-0.151	(0.61)	0.994	(1.48)
Constant	-84.023	(10.99)**	-45.197	(6.22)**

1) t-values in brackets: level of significance: * (95%), ** (99%)
2) enters as instrumental variable from step 1 (Table 5), computed for employed and non-employed women

Source: German Socio-economic Panel, wave 1, 1984; own computations,
$N_{married\ women} = 3932$, $N_{unmarried} = 2176$, $N = 12246$

Quelle: Merz 1990b

	Market wage rate (OLS)			
	married		unmarried	
	coefficient	t-value	coefficient	t-value
ln market wage rate	2.023	(0.54)[1]	1.893	(0.71)
market wage rate (DM/h)	7.56	(1.72)	6.64	(2.03)
PERSONAL CHARACTERISTICS				
Work experience				
Full-time years	0.017	(3.63)**[2]	0.044	(7.48)**
Years**2	-0.038	(3.11)**	-0.105	(6.66)**
Interruptions	0.027	(0.97)	0.038	(0.78)
Max. interruptions*10^{-3}	-0.572	(0.14)	-0.452	(0.07)
Part-time years	0.010	(1.85) *	0.017	(1.75) *
Years**2	-0.048	(2.19) *	-0.133	(3.53)**
Interruptions	-0.047	(0.82)	-0.035	(0.29)
Max. interruptions	0.007	(0.66)	0.009	(0.76)
School leaving certificate (diploma)				
Intermediate	0.112	(2.78)**	0.184	(4.07)**
Specialized upper secondary	0.036	(0.32)	0.136	(1.27)
Upper (Abitur)	0.418	(4.19)**	0.365	(4.76)**
Completed vocational training				
Apprenticeship	0.094	(2.47)**	-0.038	(0.80)
Specialized vocational	0.149	(2.85)**	0.042	(0.71)
Specialized college of higher education	0.260	(2.16) *	0.583	(3.13)**
University	0.316	(2.26) *	0.181	(1.72) *
Civil servant training	-0.164	(0.91)	-0.063	(0.34)
Health service	0.182	(2.34)**	-0.044	(0.40)
Occupational status				
Manual worker	0.823	(4.37)**	0.874	(12.58)**
White-collar worker	0.901	(4.77)**	0.996	(14.82)**
Civil servant	1.291	(5.97)**	1.069	(8.21)**
Self-employed	0.362	(1.87) *	0.639	(5.50)**
Foreigner	0.101	(2.72)**	-0.024	(0.43)
Health situation				
Impairment	-0.009	(0.14)	0.065	(0.77)
REGIONAL INFORMATION				
Community size (Boustedt)				
Up to 20.000 inhabitants	-0.117	(2.79)**	0.021	(0.36)
100.000 inhabitants and more	0.009	(0.24)	0.116	(2.24) *
Constant	0.944	(5.05)**	0.704	(10.66)
R^2	0.232		0.531	
R^2 (adj.)	0.217		0.516	

[1] Mean and standard deviation of the endogeneous variables
[2] t-values in brackets; level of significance: * (95%), ** (99%)

Source: German Socio-economic Panel, wave 1, 1984; own computations,
$N_{married\ women}$ = 3932, $N_{unmarried\ women}$ = 2176, N = 12246

Quelle: Merz 1990b

Tab. 2.4d: LIMDEP: Women's Labor Supply

In the study of the number of hours worked in a survey year by a large sample of wives, Quester and Greene (1982) were interested in whether wives whose marriages were statistically more likey to dissolve heged against that possibility by spending, on average, more time working. They reported the Tobit estimates gives in Table 22.2. The last figure in the table implies that a very large proportion of the women surveyed reported zero hours, so least squares regression would be inappropriate.

The figures in parentheses are the ratio of the coefficient estimate to the estimated asymptotic standard error. The dependent variable is hours worked in the survey year. "Small kids" is a dummy variable indicating whether there were children in the household. The "Education difference" and "Relative wage" variables compare husband and wife on these two dimensions. The wage rate used for wives was predicted using a previously estimated regression model and is thus available for all individuals, whether working or not. "Second marriage" is a dummy variable. Divorce probabilities were produced by a large microsimulation model presented in another study. The variables used here were dummy variables indicating "Mean" if the predicted probability was between 0.01 and 0.03 and "High" if it was greater than 0.03. The "slopes" are the marginal effects described earlier.

Note the marginal effects compared to the tobit coefficients. Likewise, the estimate of σ is quite misleading as an estimate of the standard deviation of hours worked.

The effects of the divorce probability variables were as expected, and were quite large. One of the questions raised in connection with this study was whether the divorce probabilities could reasonably be treated as independent variables. It might be that for these individuals, the number of hours worked was a significant determinant of the probability.

TABLE 22.2 Tobit Estimates of an Hours-Worked Equation

	White Wives		Black Wives	
	Coefficient	Slope	Coefficient	Slope
Constant	−1803.13		−2753.87	
	(−8.64)		(−9.68)	
Small kids	−1324.84	−385.89	−824.19	−376.53
	(−19.78)		(−10.14)	
Education difference	−48.08	−14.00	22.59	10.32
	(−4.77)		(1.96)	
Relative wage	312.07	90.90	286.39	130.93
	(5.71)		(3.32)	
Second marriage	175.85	51.51	25.33	11.57
	(3.47)		(0.41)	
Mean divorce probability	417.39	121.58	481.02	219.75
	(6.52)		(5.28)	
High divorce probability	670.22	195.22	578.66	264.36
	(8.40)		(5.33)	
σ	1559	618	1511	826
Sample size	7459		2798	
Proportion working	0.29		0.46	

Quelle: Greene 1993

3 Selektion und Regression - Das zweistufige Heckman Modell

3.1 Modellspezifikation, selektionskorrigierte Schätzung

Heckman 1976 hat einen zweistufigen Schätzer vorgeschlagen, der die Selektion für sich vornimmt und dann eine selektionskorrigierte Regression erlaubt; im TOBIT-Modell werden dagegen die Selektion und die Regression in einem Schritt vorgenommen. Er berücksichtigt allgemein den ‚sample selectivity bias', der auftritt, wenn Stichprobenauswahl und endogene Variable korrelieren, wenn also keine zufällige Auswahl vorliegt.

Beispiel: Positive Arbeitszeit ist Selektionskriterium und zugleich zu erklärende Variable in Arbeitsangebotsmodellen

Modell

Stufe 1: PROBIT-Schätzung für die Selektion; zugleich Schätzung der Zugehörigkeitswahrscheinlichkeit zur Auswahl aus allen Beobachtungen

$$z_i^* = \alpha' v_i + u_i$$

$$z_i = \begin{cases} 1 & falls \quad z_i > c \quad (z.B.\ c = 0) \\ 0 & sonst \end{cases}$$

mit $PROB(z_i = 1) = \Phi(\alpha' v_i / \sigma_u)$ *und* $PROB(z_i = 0) = 1 - \Phi(\alpha' v_i / \sigma_u)$

Berechnung der Mills' ratio (hazard rate)

$$\lambda_i = \phi(\alpha' v_i / \sigma_u) / \Phi(\alpha' v_i / \sigma_u) \qquad PROBIT - Ergebnis = \alpha' / \sigma_u$$

Stufe 2: OLS-Schätzung allein mit den selektierten $(z_i = 1)$ Beobachtungen aber selektionskorrigiert (COLS) mit der Selektionsvariablen λ_i

$$y_i \,|\, z_i^* > 0 = \beta' x_i + \beta_\lambda \lambda_i + \varepsilon_i$$

mit $\beta_\lambda = \rho \sigma_\varepsilon$

wobei ρ = Korrelationskoeffizient zwischen u_i und ε_i

σ_ε = Standardabweichung von ε

Dieser zweistufige Ansatz folgt aus:

$$\begin{aligned} E[y_i \,|\, x_i, selektiert] &= E y_i \,|\, x_i \\ &= E\left[y_i \,|\, x_i, \alpha' v_i + u_i > 0 \right] \\ &= \beta' x_i + E\left[\varepsilon_i / u_i > -\alpha' v_i \right] \\ &= \beta' x_i (\rho \tau_\varepsilon \sigma_u) \left\{ \Phi(-\alpha' v_i) / \left[1 - \Phi(-\alpha' v_i) \right] \right\} \\ &= \beta' x_i (\rho \tau_\varepsilon \sigma_u) \left[\Phi(\alpha' v_i) / \Phi(\alpha' v_i) \right] \end{aligned}$$

Da σ_u nicht geschätzt werden kann, wird es auf 1 normiert. Damit ist dann

$$\begin{aligned} E\left[y_i \,|\, x_i, selektiert \right] &= \beta' x_i + (\rho \sigma_\varepsilon) \lambda_i \\ &= \beta' x_i + \beta_\lambda \lambda_i \end{aligned}$$

Praktische Bedeutung:
Partizipation und Ausprägung können mit unterschiedlichen Erklärungsmustern geschätzt werden (bei TOBIT nur ein Muster).

Die korrekte asymptotische Kovarianzmatrix für den zweistufigen Schätzer ist (Heckman 1979, Greene 1995, S. 639)

$$Asy.Var\left[\hat{\beta},c\right] = \sigma_\varepsilon^2\left(x_*'x_*\right)^{-1}\left[x_*^*\left(I-\rho^2\Delta\right)x_* + \rho^2\left(x_*'\Delta V\right)\sum V'\Delta X_*\right]\left(X_*'X_*\right)^{-1}$$

$$wobei \quad X_* = \left[X : \lambda\right]$$

$$\Delta = diag\left[\delta\right]$$

$$\delta_i = -\lambda_i\left(\alpha'\vartheta_i + \lambda_i\right)\left(-1 \le \delta_i \le 0\right)$$

$$\sum = \text{asymptotische Kovarvarianzmatrix für den Schätzer von } \alpha$$

Ein konsistenter Schätzer von σ_ε^2 ist $\hat{\sigma}_\varepsilon^2 = e'e/N - \hat{\Theta}^2\overline{\delta}$.
Vorteil eines Programmpaketes wie LIMDEP: Asy.Var $\left(\hat{\beta}\right)$ wird intern berechnet und die darauf aufbauenden t-Werte für die korrekte Signifikanzbestimmung werden direkt ausgegeben.

3.2 Beispiele

- Anwendungsgebiete generell wie bei TOBIT
- Heckman's Labor Supply Model
- Labor Supply with wages and hours of work: PROBIT/COLS, Merz 1990a (Tab. 3.2a)
- Migration: Two step Heckmann approach (PROBIT/COLS) (Tab. 3.2b)
- Interdependent Multidimensional Poverty: IMD poverty risk and IMD poverty gap (Merz und Rathjen 2012

Heckman's Labor Supply Model

Heckman's (1974) labor supply model assumes that the offered wage W^o is given independently of hours of work H ():

(1) $W_i^o = \beta_2' x_{2i} + u_{2i}.$

Heckman defined the reservation wage $W_i^r = (\partial U/\partial C)/(\partial U/\partial X)$ and specified[3]

(2) $W_i^r = \gamma H_i + \alpha' z_i + v_i.$

It is assumed that the i-th individual works if

(3) $W_i^r(H_i = 0) \equiv \alpha' z_i + v_i < W_i^o$

and then the wage W_i and hours worked H_i are determined by solving (1) and (2) simultaneously after putting $W_i^o = W_i^r = W_i$. Thus we can define Heckman's model as

[3] Actually, Heckman used log W^r and W^o. The independent variables x_2 include husband's wage, asset income, prices, and individual characteristucs and z includes housewife's schooling and experience.

(4) $W_i = \beta_2' x_{2i} + u_{2i}$

and

(5) $W_i = \gamma H_i + \alpha' z_i + v_i$

for those I for which desired hours of work

(6) $H_i^* = \beta_1' x_{1i} + u_{1i} > 0,$

where $\beta_1' x_{1i} = \gamma^{-1}(\beta_2' x_{2i} - \alpha' z_i)$ and $u_{1i} = \gamma^{-1}(u_{2i} - v_i)$. Note that (3) and (6) are equivalent because $\gamma > 0$.

Call (4) and (5) the structural equations, then (3) and the identity part of (6) constitute the reduced-form equations. Heckman (1974) estimated the structural parameters by MLE (Example cited after Amemiya 1985, S. 391, 392 who shows the correspondence to a TOBIT type 3 model).

Tab. 3.2a Labor Supply with wages and hours of work: PROBIT/COLS

Table 2: Wage rate and working hours of married and unmarried women in the Federal Republic of Germany (FRG): Estimated coefficients of the PROBIT/COLS approach

	Wage rate						Labor supply (working hours)					
	married			unmarried			married			unmarried		
	coeff.	t-value	rank	coeff.	t-value	rank	coeff.	t-value	rank	coeff.	t-value	rank
PERSONAL CHARACTERISTICS												
Family status												
separate (married)	-	-	-	0.152	(1.47)	18	-	-	-	-1.544	(0.72)	37
divorced	-	-	-	0.093	(1.49)	17	-	-	-	-0.694	(0.53)	39
widowed	-	-	-	0.181	(2.00)*	16	-	-	-	1.680	(0.88)	29
Age	0.023	(0.49)	5	-0.023	(0.68)	5	-4.515	(3.51)**	2	0.898	(1.13)	2
Age *10-2	-0.041	(0.36)	3	0.129	(1.67)*	1	9.107	(2.95)**	1	-2.496	(1.35)	1
Age3*10-4	0.019	(0.21)	8	-0.139	(2.49)**	2	-7.446	(3.07)**	3	1.486	(1.23)	3
Work experience												
Full-time years	0.016	(2.58)**	6	0.012	(1.34)	8	1.288	(3.86)**	4	0.647	(2.51)**	4
Years **2	-0.034	(2.54)**	7	-0.044	(2.30)*	7	-0.650	(1.53)	12	-0.783	(1.92)*	7
Interruptions	0.028	(0.97)	22	0.015	(0.32)	32	-0.296	(0.32)	43	-0.058	(0.06)	47
Max. interruptions	0.000	(0.01)	30	-0.007	(1.02)	20	0.749	(5.37)**	9	0.292	(1.93)*	19
Part-time years	0.010	(1.29)	12	0.005	(0.42)	22	-0.171	(0.35)	19	-1.011	(2.77)**	6
Years **2	-0.045	(1.88)*	11	-0.092	(2.42)**	11	1.233	(1.15)	14	4.277	(5.07)**	5
Interruptions	-0.048	(0.85)	21	-0.044	(0.37)	29	1.006	(0.60)	36	-3.992	(1.66)*	23
Max. interruptions	0.006	(0.65)	24	0.009	(0.74)	23	-0.222	(0.77)	35	0.085	(0.33)	40
Vocational and further training	-0.003	(0.02)	29	-0.042	(0.59)	24	-5.955	(1.73)*	27	-4.373	(2.19)*	13
School-leaving certificate (diploma)												
Intermediate	0.119	(2.92)**	16	0.166	(3.42)**	12	1.384	(1.06)	31	-0.014	(0.01)	49
Specialized upper secondary	0.044	(0.39)	26	0.148	(1.42)	19	0.202	(0.06)	50	0.167	(0.08)	46
Upper (Abitur)	0.427	(4.26)**	9	0.337	(4.30)**	10	-2.800	(1.00)	30	-5.042	(3.13)**	14
Completed vocational training												
Apprenticeship	0.090	(2.39)**	18	-0.051	(1.09)	21	-0.109	(0.10)	47	1.888	(1.85)*	21
Specialized vocational	0.149	(2.87)**	17	0.023	(0.41)	31	-0.732	(0.51)	39	1.359	(1.11)	30
Specialized college of higher education	0.259	(2.16)*	19	0.592	(3.25)**	14	6.984	(2.00)*	22	-4.624	(1.26)	31
University	0.308	(2.20)*	13	0.070	(0.64)	25	3.728	(0.97)	28	8.385	(3.22)**	12
Civil servant training	-0.147	(0.82)	23	-0.109	(0.59)	27	0.460	(0.09)	48	3.332	(0.90)	33
Health service	0.183	(2.35)**	20	-0.054	(0.50)	30	6.916	(2.91)**	18	2.953	(1.35)	28
Occupational Status												
Manual worker	0.798	(3.56)**	2	0.808	(9.51)**	4	-7.288	(1.49)	8	-5.092	(3.04)**	10
White-collar worker	0.877	(3.92)**	1	0.928	(11.29)**	3	-5.806	(1.19)	10	-4.999	(3.12)**	9
Civil servant	1.258	(5.04)**	4	1.029	(7.96)**	6	-4.259	(0.77)	26	-5.149	(2.05)*	16
Self-employed	0.334	(1.46)	10	0.655	(5.20)**	9	1.995	(0.39)	33	-0.552	(0.22)	43
Foreigner	0.100	(2.66)**	15	-0.027	(0.52)	28	8.022	(6.31)**	7	-0.073	(0.07)	48
Health Situation												
Impairment 1)	-	-	-	-	-	-	0.948	(1.29)	29	-0.527	(0.64)	36
Chronic complaints 2)	-	-	-	-	-	-	-1.04	(0.97)	34	0.630	(0.59)	38

FS Tab. 3.2a Labor Supply with wages and hours of work: PROBIT/COLS

Certified handicapped 2)	-0.010	(0.15)	28	0.065	(0.78)	26	0.430	(0.22)	45	-2.048	(1.10)	32
(Subjective) satisfaction 3)												
Household income	-	-	-	-	-	-	0.555	(3.19)**	16	0.336	(1.88)*	22
Household activities	-	-	-	-	-	-	-0.927	(4.54)**	13	-0.455	(2.69)**	18
Monthly unearned income												
Old age, widow's pension 6)	-	-	-	-	-	-	-16.896	(1.90)*	24	-12.433	(3.95)**	8
Private support 5)6)	-	-	-	-	-	-	-19.592	(1.36)	32	-9.421	(3.04)**	15
Unemployment benefit/relief (t-1)	-	-	-	-	-	-	-7.673	(2.19)*	23	-5.011	(1.29)	24
HUSBAND'S CHARACTERISTICS												
Occupational status												
Manual Worker	-	-	-	-	-	-	1.960	(1.25)	20	-	-	-
White-collar worker	-	-	-	-	-	-	0.061	(0.03)	51	-	-	-
Civil servant	-	-	-	-	-	-	-1.012	(0.44)	38	-	-	-
Self-employed	-	-	-	-	-	-	9.165	(4.56)**	11	-	-	-
Net earned income last month 6)	-	-	-	-	-	-	-0.638	(1.47)	25	-	-	-
HOUSEHOLD CHARACTERISTICS												
Nursing cases existing 2)	-	-	-	-	-	-	0.494	(0.18)	46	3.310	(1.61)	26
Children in age classes												
0 to 5 years	-	-	-	-	-	-	-8.341	(3.98)**	6	-2.554	(1.41)	27
6 to 9 years	-	-	-	-	-	-	-2.913	(2.73)**	17	0.408	(0.26)	44
10 to 15 years	-	-	-	-	-	-	-2.275	(3.26)**	15	-0.160	(0.22)	45
16 and older	-	-	-	-	-	-	-0.233	(0.44)	41	-0.122	(0.33)	42
Community size (Bousted)												
up to 20.000 inhabitants	-0.116	(2.79)**	14	0.011	(0.19)	33	0.654	(0.56)	37	-0.828	(0.72)	35
100.000 inhabitants and more	0.011	(0.29)	27	0.123	(2.42)**	13	0.079	(0.08)	49	-1.880	(1.76)*	20
Monthly unearned income: transfer payments												
Housing benefits 6)	-	-	-	-	-	-	-6.425	(0.46)	42	-44.271	(2.61)**	17
Welfare benefits 5)6)	-	-	-	-	-	-	6.767	(0.48)	44	-2.665	(0.33)	41
Yearly unearned income: property												
Income from let and lease 6)	-	-	-	-	-	-	-0.377	(2.61)**	21	0.201	(1.68)*	25
Income from interests and dividends 6)	-	-	-	-	-	-	0.095	(0.51)	40	-0.414	(1.03)	34
Constant	0.573	(0.93)	31	0.964	(2.33)*	34	95.463	(5.62)**	52	31.936	(2.51)**	50
Lambda	0.026	(0.42)	25	-0.143	(1.46)	15	14.091	(2.80)**	5	5.557	(1.07)	11
R2(adj.)	0.22	-	-	0.54	-	-	0.37	-	-	0.24	-	-

n (wage rate) weekly working hours
evel of significance: * (95%), ** (99%); rank is due to absolute standardized B-coefficients, t-values due to the correct asymptotic variance covariance matrix.
) 0 = not at all, 1 = slightly, 2 = considerably 2) 0 = no, 1 = yes 3) scale 0-10, 0 = completely disconten, 10 = completely content 4) from private persons not living in the household 5) inclusive subsistence
benefits/benefits granted in needy situations 6) *10-3;
Source: Socio-Economic Panel 1984, Wave 1, own computations

Quelle: Merz 1990a, S. 256, 257, PROBIT: siehe Abschnitt II.6.4

Migration: Two step Heckmann approach (PROBIT/COLS)

Nakosteen and Zimmer (1980) analyze a model of migration based on the following structure:[5]

For individual, i, the market wage they can earn at their present location is

$$y_p^* = \beta' x_p + \varepsilon_p.$$

Variables in the equation include age, sex, race, growth in employment, and growth in per capita income. If the individual migrates to a new location, his market wage would be

$$y_m^* = \gamma' x_m + \varepsilon_m$$

[4]Unless there is some compelling reason, binomial probability models should not be estimated without constant terms.

[5]A number of other studies have also used variants of this basic formulation. Some important examples are Willis and Rosen (1979) and Robinson and Tomes (1982). The study by Tunali examined in Example 21.8 is another example.

However, migration entails costs that are related both to the individual and to the labor market,

$$C^* = \alpha' z + u.$$

Costs of moving are related to whether the individual is self-employed and whether the individual recently changed their industry of employment. The individual migrates if the benefit, $y_m^* - y_p^*$, is greater than the cost, C^*. The net benefit of moving is

$$\begin{aligned} M^* &= y_m^* - y_p^* - C^* \\ &= \gamma' x_m - \beta' x_p - \alpha' z + (\varepsilon_m - \varepsilon_p - u) \\ &= \delta' w + \varepsilon. \end{aligned}$$

We could use an ordinary regression but for the fact that M^* is unobservable. The individual either moves or does not. After the fact, we observe only y_m^* if the individual has moved or y_p^* if he or she has not. But we do observe that $M = 1$ for a move and $M = 0$ for no move. If the disturbances are normally distributed, this produces the probit model we analyzed earlier. Logistic disturbances produce the logit model instead.

Example 21.3 presents a model of migration analyzed by Nakosteen and Zimmer. That model fits precisely into the framework described earlier. Briefly, the equations of the model are

Net benefit of moving:	$M_i^* = \gamma' w_i + u_i,$
Income if moves:	$I_{i1} = \beta_1' x_{i1} + \varepsilon_{i1},$
Income if stays:	$I_{i0} = \beta_0' x_{i0} + \varepsilon_{i0}.$

One of the components of the net benefit is the market wage individuals could achieve if they move, compared to what they could obtain if they stay. Therefore, among the determinants of the net benefit are factors that also affect the income received in either place.

An analysis of income in a sample of migrants must account for the incidental truncation of the mover's income on a positive net benefit. Likewise, the income of the stayer is incidentally truncated on a nonpositive net benefit. The model implies an income after moving for all observations, but we observe it only for those who actually do move.

FS Migration: Two step, Heckmann approach (PROBIT/COLS)

Nakosteen and Zimmer (1980) applied the model of Example 22.16 to a sample of 9223 individuals with data for 2 years (1971 and 1973) sampled from the Social Security Administration's Continuous Work History Sample. Over the period, 1078 individuals migrated and the remaining 8145 did not. The independent variables in the migration equation were as follows:

SE = self-employment dummy variable; 1 if yes,

ΔEMP = rate of growth of state employment,

ΔPCI = growth of state per capita income,

AGE, RACE (nonwhite = 1), SEX (female = 1),

ΔSIC = 1 if individual changes industry.

The earnings equations included Δ*SIC* and *SE*. The authors reported the results given in Table 22.5. The figures in parentheses are asymptotic *t* ratios.

TABLE 22.5

	Migration	Migrant Earnings	Nonmigrant Earnings
Constant	−1.509	9.041	8.593
SE	−0.708	−4.104	−4.161
	(−5.72)	(−9.54)	(−57.71)
ΔEMP	−1.488	—	—
	(−2.60)		
ΔPCI	1.455	—	—
	(3.14)		
AGE	−0.008	—	—
	(−5.29)		
RACE	−0.065	—	—
	(−1.17)		
SEX	−0.082	—	—
	(−2.14)		
ΔSIC	0.948	−0.790	−0.927
	(24.15)	(−2.24)	(−9.35)
λ	—	0.212	.863
		(0.50)	(2.84)

Quelle: Greene 1993

Interdependent Multidimensional Poverty:
IMD poverty risk and IMD poverty gap (Merz und Rathjen 2012)

Merz, J. and T. Rathjen (2012), *Intensity of Time and Income Interdependent Multidimensional Poverty: Well-Being and Minimum 2DGAP – German Evidence*, Forschungsinstitut Freie Berufe (FFB), Leuphana Universität Lüneburg, Revised FFB-Discussion Paper No.92 (2011), Lüneburg.

... What drives Time and Income Uni- and Multidimensional Poverty – Poverty Risk and Poverty Gap Microeconometric Estimation

Two important questions emerge after the above description of uni- and multidimensional time and income poverty in Germany: First, what drives the IMD poverty risk and the minimum poverty 2DGAP, and second, have uni- and multidimensional time and income poverty risk and poverty gap to be differently explained on the individual level.

Since the respective poor are expected not to be a random subgroup of the entire population, respectively of the working population, we apply the Heckman (1976) approach to account for the expected selectivity. As known, the two step procedure with step 1: PROBIT estimation for the probability to be IMD poor and step 2: selectivity correction estimates of the minimum 2DGAP allows a separate explanation of the probability of being selected, the poverty risk, and the amount of the variable of interest, the poverty gap.

gap_näe: IMD poverty gap (2DGAP):
incomeheck: 0,1 Dummy if IMD poor
Daten: Zeitbudgeterhebung 2001/02

```
. /* IMD poverty gap (2DGAP) heckman only workers step 1=f(...humcap...)*/
. heckman gap_näe $personal $humcap $occupation2 $job $household $hh ost, ///
> select (incomeheck = frau alter alter_2 nicht_deutsch $education1 $occupation2
$job $household $branch ost) vce(cluster idhh)

Iteration 0:   log pseudolikelihood = -2890.8785
Iteration 1:   log pseudolikelihood = -2885.6395
Iteration 2:   log pseudolikelihood = -2885.2314
Iteration 3:   log pseudolikelihood =  -2885.208
Iteration 4:   log pseudolikelihood = -2885.2076
Iteration 5:   log pseudolikelihood = -2885.2076

Heckman selection model                         Number of obs      =      7354
(regression model with sample selection)        Censored obs       =      7039
                                                Uncensored obs     =       315

                                                Wald chi2(31)      =     67.48
Log pseudolikelihood = -2885.208                Prob > chi2        =    0.0002

                              (Std. Err. adjusted for 2877 clusters in idhh)
------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
gap_näe      |
        frau |   41.73643   29.73742     1.40   0.160    -16.54784    100.0207
       alter |  -35.21333   123.4609    -0.29   0.775    -277.1923    206.7656
     alter_2 |   .2875511   .2924849     0.98   0.326    -.2857087    .8608109
     married |   51.47017   26.32988     1.95   0.051    -.1354504    103.0758
       phact |  -.9432926   1.053984    -0.89   0.371    -3.009064    1.122479
nicht_deut~h |   80.07997    52.0209     1.54   0.124    -21.87911    182.0391
    yrschool |   19.44577   119.0233     0.16   0.870    -213.8356    252.7272
    experien |   20.59087   121.3278     0.17   0.865    -217.2072     258.389
    experie2 |   -.183953   .2826168    -0.65   0.515    -.7378717    .3699657
```

```
freiberufler |   41.81865   64.87999     0.64   0.519    -85.34379    168.9811
 unternehmer |   104.6262   63.48152     1.65   0.099    -19.79527    229.0477
    arbeiter |   48.10112   56.86632     0.85   0.398    -63.35482    159.5571
angestellter |   50.55604   49.63349     1.02   0.308     -46.7238    147.8359
anderer_be~f |   71.25721   62.52065     1.14   0.254    -51.28102    193.7954
        wage |   .8374102   7.007458     0.12   0.905    -12.89695    14.57178
wtl_arbeit~h |  -.6614543   1.304317    -0.51   0.612    -3.217868     1.89496
        cat2 |  -.6438433   16.65603    -0.04   0.969    -33.28907    32.00138
        cat3 |  -28.68344   61.83363    -0.46   0.643    -149.8751    92.50824
        cat4 |   21.81896   48.42358     0.45   0.652    -73.08952    116.7274
    couple_1 |  -15.01466   49.97334    -0.30   0.764    -112.9606    82.93129
    couple_2 |   40.35428   47.26357     0.85   0.393    -52.28062    132.9892
couple_mehr2 |  -21.11259     60.494    -0.35   0.727    -139.6786    97.45347
      single |    76.5522   48.65279     1.57   0.116    -18.80552    171.9099
    single_1 |  -2.524153   57.64996    -0.04   0.965     -115.516    110.4677
single_mehr1 |   87.32596   69.01956     1.27   0.206     -47.9499    222.6018
  anderer_hh |   60.34263   70.95754     0.85   0.395    -78.73158    199.4168
    szeithhh |   1.372925   6.016606     0.23   0.819    -10.41941    13.16526
    zeitkinh |  -34.96635    12.5792    -2.78   0.005    -59.62113   -10.31157
    h_jungki |   58.84742   30.10268     1.95   0.051    -.1527571    117.8476
       hhpas |   1.467475   1.394547     1.05   0.293    -1.265787    4.200737
         ost |  -3.023493   22.86664    -0.13   0.895    -47.84128    41.79429
       _cons |   477.0374   836.8789     0.57   0.569    -1163.215     2117.29
-------------+----------------------------------------------------------------
incomeheck   |
        frau |  -.2510278   .1053766    -2.38   0.017    -.4575622   -.0444935
       alter |   .0912819   .0347081     2.63   0.009     .0232552    .1593087
     alter_2 |  -.0010285   .0004348    -2.37   0.018    -.0018806   -.0001763
nicht_deut~h |   .9243942   .2783825     3.32   0.001     .3787745    1.470014
  realschule |  -.0242668   .1158695    -0.21   0.834    -.2513669    .2028332
 hauptschule |  -.0846072   .1338442    -0.63   0.527     -.346937    .1777225
 noch_schule |  -5.221337   .2746846   -19.01   0.000    -5.759709   -4.682965
freiberufler |   .9270949   .3195762     2.90   0.004     .3007371    1.553453
 unternehmer |    .945742   .3225095     2.93   0.003     .3136351    1.577849
    arbeiter |   .5398324   .2822636     1.91   0.056     -.013394    1.093059
angestellter |    .261927   .2778326     0.94   0.346    -.2826148    .8064688
anderer_be~f |   .4083565   .4036332     1.01   0.312      -.38275    1.199463
        wage |  -.1523297   .0368902    -4.13   0.000    -.2246332   -.0800263
wtl_arbeit~h |  -.0202908   .0049587    -4.09   0.000    -.0300096   -.0105719
        cat2 |   .1421552   .0893566     1.59   0.112    -.0329806     .317291
        cat3 |  -.0310595    .186515    -0.17   0.868    -.3966221    .3345031
        cat4 |  -.0735581    .261323    -0.28   0.778    -.5857418    .4386256
    couple_1 |   .0784435   .1889899     0.42   0.678    -.2919699     .448857
    couple_2 |    .204379   .1871951     1.09   0.275    -.1625167    .5712747
couple_mehr2 |   .9412708   .2159297     4.36   0.000     .5180563    1.364485
      single |   .6074646   .1991268     3.05   0.002     .2171833    .9977458
    single_1 |   1.110563   .2015303     5.51   0.000      .715571    1.505555
single_mehr1 |   1.179567   .2296642     5.14   0.000      .729433      1.6297
  anderer_hh |   .3184429   .2745629     1.16   0.246    -.2196905    .8565764
    agricult |   .3834735   .2135671     1.80   0.073    -.0351102    .8020573
    industry |   .0817127   .1129351     0.72   0.469     -.139636    .3030613
         ost |   .1635432   .1117223     1.46   0.143    -.0554285    .3825148
       _cons |  -2.529793   .8130886    -3.11   0.002    -4.123417   -.9361688
-------------+----------------------------------------------------------------
     /athrho |  -.4533076   .3223699    -1.41   0.160    -1.085141    .1785257
    /lnsigma |   4.744404   .1461659    32.46   0.000     4.457924    5.030884
-------------+----------------------------------------------------------------
         rho |  -.4246141   .2642475                     -.7950978     .176653
       sigma |   114.9393    16.8002                     86.30818    153.0683
      lambda |  -48.80485   36.88258                    -121.0934    23.48367
------------------------------------------------------------------------------
Wald test of indep. eqns. (rho = 0): chi2(1) =     1.98   Prob > chi2 = 0.1597
------------------------------------------------------------------------------
```

```
. eststo cfinincome
```

4 Selektion und Regression – Das generalisierte zweistufige Modell von Lee

4.1 Struktur und Schätzung des Ansatzes

Lee (1983) hat den zweistufigen Heckman-Ansatz verallgemeinert und erlaubt eine generellere Spezifikation der Selektionsgleichung

Idee

Berechnung der Auswahlwahrscheinlichkeit nicht durch PROBIT, sondern durch eine andere Spezifikation (z.B. LOGIT). Diese Wahrscheinlichkeit bestimmt dann die Stelle an der die Standardnormalverteilung für die Mills' ratio berechnet wird.

Modell

Stufe 1: Selektionswahrscheinlichkeit (z.B. durch LOGIT)

$$z_i^* = \alpha' \vartheta_i + \mu_i$$

$$z_i = \begin{cases} 1 \;\; falls \;\; z_i^* > c & (z.B. \;\; c = 0) \\ 0 \;\; sonst \end{cases}$$

Transformation (für das LOGIT-Modell)

$$q = \Phi^{-1}\left[e^{\alpha' \vartheta} / \left(1 + e^{\alpha' \vartheta}\right) \right] = \Phi^{-1}\left[P_{LOGIT} \right]$$

Mills' ratio

$$\lambda = \phi(q) / \Phi(q)$$

Stufe 2: wie bei Heckman

$$y_i \left| z_i^* > 0 = \beta' x_i + \beta_\lambda \lambda_i + \varepsilon_i \right.$$

Ist bspw. die Selektion komplexer, z.B. Auswahl durch ein MULTINOMIALES LOGIT-Modell, dann ist die Berechnung vor allem der korrekten asymptotischen Varianz-Kovarianzmatrix von β, β_λ rel. aufwendig (vgl. Greene 1995, S. 649 und Beispiel (b)).

4.2 Beispiele Lee Ansatz

- Markt- und nichtmarktmäßige Aktivitäten privater Haushalte: Partizipation in Haupterwerb (HET), Nebenerwerb (NET), Eigenarbeit (EA) und Kombination dieser Aktivitäten (MULTINOMIALES LOGIT (8 Alternativen/COLS); Merz 2012, S. 335) (Tab: 4.2a)
- LIMDEP: MNL/COLS-Program (Tab: 4.2b

Tab. 4.2a Markt- und nichtmarktmäßige Aktivitäten privater Haushalte (Merz 2012)

D.6.1.1 AKTIVITÄTSKOMBINATIONEN UND NUTZENTHEORETISCHE FUNDIERUNG DER PARTIZIPATIONSENTSCHEIDUNG UND DER AKTIVITÄTS-/ARBEITSZEITEN IM MNL/COLS-MODELL

Wie eben schon angesprochen, erlaubt unsere Mikrodatenbasis die Analyse der Aktivitäten Haupterwerb (H), Nebenerwerb (N) und Eigenarbeit (E). Abbildung AD.2 benennt die einzelnen Kombinationen oder *Alternativen* j (j=1,...,m) aller m=8 Möglichkeiten.

Abbildung AD.2: Mögliche Kombinationen (Alternativen j) und zugeordnete Nutzen u_j markt- und nichtmarktmäßiger Aktivitäten: Haupterwerb (H), Nebenerwerb (N) und Eigenarbeit (E) sowie Freizeit (F)

j	H	N	E	F	Nutzen
1	1	0	0	1	$u_1(t_H, 0, 0, t_F)$
2	0	1	0	1	$u_2(0, t_N, 0, t_F)$
3	0	0	1	1	$u_3(0, 0, t_E, t_F)$
4	1	1	0	1	$u_4(t_H, t_N, 0, t_F)$
5	1	0	1	1	$u_5(t_H, 0, t_E, t_F)$
6	0	1	1	1	$u_6(0, t_N, t_E, t_F)$
7	1	1	1	1	$u_7(t_H, t_N, t_E, t_F)$
8	0	0	0	1	$u_8(0, 0, 0, t_F)$

Mit der Abbildung AD.2 kennzeichnen zunächst die ersten drei Alternativen (j=1,2,3) jeweils ausschließliche Aktivitäten, die nächsten drei Alternativen (j=4,5,6) Zweierkombinationen, Alternative j=7 Aktivitäten sowohl im Haupt-, Nebenerwerb als auch in der Eigenarbeit und schließlich Alternative j=8 als die Alternative, die ohne diese Tätigkeiten 'reine' Freizeit (F) ist.

Jeder Alternative j (j=1,...,8) sei ein Nutzen u_j zugeordnet, der von den Ausprägungen der jeweiligen Aktivitäten y_l abhängt. Wir untersuchen die Aktivitäts-/Arbeitszeit der Aktivitäten Haupterwerb, Nebenerwerb und Eigenarbeit (und Freizeit als verbleibende Kategorie), so daß die jeweilige Zeit t_l der Ausprägung y_l (l=H,N,E) entspricht (vgl. Abbildung AD.2). Der Nutzen jeder weiteren Alternative u_j ist eine Funktion exogener und zufälliger Variablen und entspreche dem Nutzenkonzept des im Abschnitt B.2 entwickelten Modells multipler markt- und nichtmarktmäßiger Aktivitäten der Stufe 3 (Inputraum) aus Gleichung (B.214b), wenn von Marktgüterinputs abgesehen wird.

Für jede Alternative j wird angenommen, daß die Nutzenmaximierung unter der *Zeitrestriktion*

$$t = t_H + t_N + t_E + t_F,$$

mit t als die insgesamt zur Verfügung stehende und aufzuteilende Zeit, eine innere Lösung hervorbringt. Die ***nutzenmaximierenden Aktivitätszeiten*** seien

$$(t_{1H}^*), \quad (t_{2N}^*), \quad (t_{3E}^*), \quad (t_{4H}^*, t_{4N}^*), \quad (t_{5H}^*, t_{5E}^*),$$
$$(t_{6N}^*, t_{6E}^*) \text{ und } (t_{7H}^*, t_{7N}^*, t_{7E}^*)$$

mit entsprechend verbleibender Zeit t_F^* der jeweiligen Alternative. Diese nutzenmaximalen Zeiten, die eine Funktion der nutzenexogenen und nutzenzufälligen Variablen und der Gesamtzeit sind, seien durch lineare Funktionen approximiert

(D.56) $t_{jl}^* = \beta_{jl}'x + \varepsilon_{jl} \qquad (j=1,\ldots,8),(l=H,N,E),$

wobei x ein Vektor erklärender, sozioökonomischer Variablen, β_{jl} der in der zweiten Stufe (COLS-Teil) zu bestimmende Parametervektor und ε_{jk} ein entsprechender Fehlerterm ist.

Mit $v_j^* = v_j^*(t_{jk}^*)$ $(j=1,\ldots,7)$ seien die *indirekten Nutzen* bei nutzenmaximalen Aktivitätszeiten t_{jk}^* beschrieben; für die reine Freizeitalternative j=8 gelte $v_8 = u(0,0,0,t)$.[1)]

Dubin und McFadden 1984, S.346-350 wählen eine spezifische indirekte Nutzenfunktion, aus der sie ihren multinomialen Ansatz ableiten. Wir approximieren (wie auch Amemiya et al. 1987) die indirekten Nutzenfunktionen linear als

(D.57) $u_j = v_j^* = v_j + \varepsilon_j = \alpha_j'x + \varepsilon_j,$

mit den gesuchten Parametern α_j der alternativenspezifischen Determinanten.

Eine Alternative j wird schließlich gewählt, wenn $u_j > u_k$ $(j,k=1,\ldots,8;\ j\neq k)$. Mit der Annahme, daß ε_j alternativenunabhängig extremwertverteilt ist, folgt die Auswahl dann dem in Abschnitt D.3.2 beschriebenen MULTINOMIALEN LOGIT (MNL)-Modell.

Tabelle D.1: Markt- und nichtmarktmäßige Aktivitäten: Partizipation in Haupterwerb (HET), Nebenerwerb (NET), Eigenarbeit (EA) und Kombinationen dieser Aktivitäten. Diskretes Entscheidungsmodell: Multinomialer Logit (8 Alternativen)

	Eine Tätigkeit			Zwei Tätigkeiten			Drei Tätigkeiten
Partizipation	HET	NET	EA	HET ∩ NET	HET ∩ EA	NET ∩ EA	HET ∩ NET ∩ EA
j[1]	1	2	3	4	5	6	7
Stichprobenanteil [%]	25.9	3.0	10.6	1.6	9.8	2.2	2.0
PERSÖNLICHE MERKMALE							
Geschlecht							
Frau	-1.484**[2]	-0.352**	0.529**	-1.462**	-1.157**	-0.060	-1.765**
Familienstand							
verheiratet	0.601**	-0.656**	0.378**	0.856**	0.929**	0.177	0.304
Alter	-0.071**	-0.043**	-0.011**	-0.075**	-0.075**	-0.052**	-0.072**
Ausbildungsabschluß							
+Haupt-/Realschule[3]							
Abitur/Hochschulabschluß	-0.746**	0.398*	0.162	-1.019**	-0.669**	0.847**	-1.787**
Soziales Netzwerk	0.240**	0.822**	1.365**	0.583**	1.593**	1.546**	1.521**
HAUSHALTSMERKMALE							
Haushaltsgröße	-0.491**	0.153*	0.007	-0.169*	-0.509**	0.129	-0.266**
Kinder unter 6 Jahren[4]	-0.137	-0.833**	0.031	-0.636*	-0.377**	-0.081	-0.425
Haushaltsnettoeinkommen (Monat)							
+< 1500							
≥ 1500	1.896**	0.223	0.244*	0.913**	1.764**	-0.280	1.918**
SAISON							
+Herbst/Winter							
Frühjahr/Sommer	-0.019	0.164	0.101	0.282	0.356**	0.065	0.712**
REGIONAL- UND ARBEITSMARKTINFORMATIONEN							
Dauerarbeitslosenquote	-0.009**	-0.014**	-0.010**	-0.025	-0.013**	-0.017**	-0.037**
Konstante	3.031**	-0.667**	-1.916**	0.602	1.385**	-1.189**	0.108
Log(likelihood)	-8856.0						
Pseudo R^2	38.0%						

1) j kennzeichnet die jeweilige Alternative (Aktivitätskombination), Normalbasierung j=8 (nur Freizeit 44.9% aller Befragten)
2) Signifikanzniveau: * (95%), ** (99%)
3) +: Referenzkategorie
4) 1-Kinder in der entsprechenden Altersklasse, 0-sonst.

Quelle: Sfb 3 - Nebenerwerbstätigkeitsumfrage 1984, N=7098, eigene Berechnungen

Tabelle D.2a: Markt- und nichtmarktmäßige Aktivitäten: Kombinationen wöchentlicher Arbeitszeit in Haupterwerb (HET), Nebenerwerb (NET) und Eigenarbeit (EA). Generalisiertes diskret/kontinuierliches Entscheidungsmodell: Multinomiale selektionskorrigierte Schätzung der Gesamtarbeitszeit: MNL/COLS.

	Eine Tätigkeit			Zwei Tätigkeiten			Drei Tätigkeiten
	HET	NET	EA	HET ∩ NET	HET ∩ EA	NET ∩ EA	HET ∩ NET ∩ EA
Gesamte ø wöchentliche Arbeitszeit[1]	40.43	7.06	5.86	42.76	43.13	9.82	46.42
Standardabweichung	(8.28)	(8.53)	(9.22)	(11.85)	(10.96)	(6.98)	(10.29)
N	1833	212	755	115	696	154	144
j[2]	1	2	3	4	5	6	7
PERSÖNLICHE MERKMALE							
Geschlecht							
Frau	-1.753[3]	2.158	0.220	-6.280	-3.130	-1.172	-9.160
Familienstand							
verheiratet	-3.045**	-0.737	0.413	-3.044	-7.883**	3.801*	-2.646
Alter							
Alter	0.395**	0.125	0.101	1.248*	-0.043	-0.403	-0.130
Alter$^2 \cdot 10^{-2}$	-0.266**	-0.124	-0.140	-1.586*	0.490*	0.435	0.335
Ausbildungsabschluß							
+Ohne Abschluß/Hauptschule[4]							
Realschule mit Lehre	-0.754*	-1.820	-0.231	-2.149	1.215	-0.011	-0.191
Abitur	-0.750	-0.671	0.400	-0.483	-0.927	3.248	12.483
Hochschulabschluß	3.662**	-0.982	-3.090*	-8.626	-0.173	1.877	0.854
Berufliche Stellung							
+Berufsausbildung/Schüler/Student							
Arbeiter	0.273	-	-	-5.353	1.170	-	-1.665
Angestellter	1.733	-	-	2.207	2.008	-	-0.104
Beamter	1.726	-	-	-3.513	1.650	-	8.528*
Selbständiger	13.764**	-	-	9.588	11.186**	-	6.364
Erwerbsloser	-	1.237	0.674	-	-	-2.820	-
Rentner	-	-0.426	2.479	-	-	-8.117*	-
Haushaltsführender	-	-4.492	1.330	-	-	-2.126	-
Soziale Selbsteinstufung							
+Unterschicht							
Mittelschicht	-0.337	1.725	0.294	-1.600	-2.059	-0.541	0.311
Oberschicht	0.850	2.486	1.777	0.043	-2.273	-1.885	-3.194
AKTIVITÄTSMERKMALE							
Nebenerwerbstätigkeit							
Aktivitäten							
+Landwirtschaft							
Produzierendes Gewerbe	-	-0.938	-	-5.889	-	-2.563	0.379
Dienstleistung	-	2.225	-	-8.046	-	-2.185	1.575
Schwarzarbeit	-	3.916**	-	0.479	-	2.031	1.585
Eigenarbeit (DIY)							
Aktivitäten							
Handarbeit	-	-	4.445**	-	3.824**	6.385**	7.726*
Renovieren	-	-	2.295*	-	1.719*	3.063*	2.262
Hausbau	-	-	2.750	-	5.701**	0.787	6.072*
Autoreparatur	-	-	1.226	-	0.054	-0.300	1.197
Gartenarbeit	-	-	4.109**	-	6.326**	4.317*	4.910*
Kunsthandwerk	-	-	1.393	-	3.879**	3.444	3.370
+Sonstiges							

Fortsetzung Tabelle D.2a:

	Eine Tätigkeit			Zwei Tätigkeiten			Drei Tätigkeiten
	HET	NET	EA	HET ∩ NET	HET ∩ EA	NET ∩ EA	HET ∩ NET ∩ EA
j	1	2	3	4	5	6	7
Soziales Netzwerk	0.888*	-1.068	0.673	-5.014*	-7.927**	6.473**	-0.910
Lohnsatz							
Haupterwerb	-0.394**	-	-	-0.279*	-0.439**	-	-0.471**
Nebenerwerb	-	-0.114**	-	-0.038	-	-0.001	-0.129*
Eigenarbeit	-	-	-0.125**	-	-0.016	-0.032	-0.036
HAUSHALTSMERKMALE							
Haushaltsgröße	1.051**	-0.348	-0.594	-1.148	2.417*	1.105	-0.281
Kinder[5]							
0 - 2 Jahre	0.073	-4.970	3.277*	-0.506	1.862	1.665	1.737
3 - 5 Jahre	0.472	7.737**	0.005	-1.231	1.399	-2.728	4.471
Haushaltsnettoeinkommen (Monat)							
+<1500							
1500 - <2500	1.037	-0.801	-0.699	2.456	5.263**	0.816	4.422
≥2500	2.123**	-1.323	1.100	3.218	5.670**	-0.975	4.128
SAISON							
Frühjahr	-5.936**	-1.137	-0.702	-0.703	-8.640**	-5.651*	-4.909
Sommer	-5.957**	-0.803	-0.334	-5.368	-8.774**	-6.184*	-1.422
Herbst	-1.152**	0.229	-0.741	-5.536	-2.674*	-0.809	1.881
+Winter							
REGIONAL- UND ARBEITSMARKTINFORMATIONEN							
Ortsgröße <5000 Einwohner	0.688	2.328	-0.819	1.465	-0.984	-0.941	-1.925
Beschäftigte im tertiären Sektor/Alle	-0.008	0.030	-0.052	-0.212	0.013	0.055	-0.333*
Dauerarbeitslosenquote	0.006	-0.041	-0.016	0.016	0.005	-0.083	0.032
Bruttowertschöpfung/Einwohner 10^{-3}	-0.037	0.107	-0.013	0.148	-0.113	-0.093	0.185
Lambda[6]	-9.102**	3.556	2.860	-7.994	-16.159**	12.600*	-6.562
Konstante	45.355**	-7.287	0.283	71.743**	78.178**	-8.758	81.430
R^2	32.5%	21.7%	12.6%	36.9%	21.7%	29.6%	40.1%
R^2 adj.	31.4%	8.8%	8.6%	12.3%	17.5%	7.2%	17.6%

1) Gesamte Arbeitszeiten in den jeweils ausgeübten Tätigkeiten
2) j kennzeichnet die jeweilige Alternative (Aktivitätskombination), Normalbasierung j=8 (nur Freizeit)
3) Signifikanzniveau: * (95%), ** (99%)
4) +: Referenzkategorie
5) 1-Kinder in der entsprechenden Altersklasse, 0-sonst.
6) Selektionskorrekturvariable

Quelle: Merz 2012

Tab. 4.2b LIMDEP: MNL/COLS-Program

```
================================================================
Sample selection with multinomial logit selection model         =
Requires extensive modification by user for application.        =
Note lines below marked by *                                    =
                                                                =
User must define the data matrix with the following line:       =
                                                                =
        NAMELIST         ; V = RHS in logit model               =
                                                                =
 User's dependent variable is Y.  Change the following line     =
                                                                =
        CREATE           ; Y = the dependent variable $         =
                                                                =
Logit dependent variable is Z.  Change the following line       =
                                                                =
        CREATE           ; Z = the logit dependent variable $   =
================================================================

Estimate the multinomial logit model.  User supplies the name of z.

       LOGIT          ; LHS = z  ; RHS = V $

Pick off the subvectors of à in blocks of K2.  Where the values taken by z are
0,1,...,J, there are J blocks of K2 coefficients. User must modify this to put the
right values in the PART commands. We also recover the covariance matrix, SG.

(*)    MATRIX         ; a1 =  part(b,1,K2) ; a2 = part(b,...)  ; ...
                      ; SG  =  varb $

Compute the predicted probabilities for the logit model.

       CREATE  ; p1 = exp(dot(V,a1))
               ; p2 = exp(dot(V,a2))
(*)            ; ... (do this J times)
               ; p0 = 1 / (1 + p1 + p2 + ...)
               ; p1 = p1 * p0 ; p2 = p2 * p0 ; ... $

Do the sample selection.

(*)    INCLUDE ; NEW ;  z = ...  $ (your value to select on)

Compute lambda and delta

(*)    CREATE  ; H = inp(...)    ? (insert pj = the prob. of your choice)
               ; lambda = n01(H) / phi(H)
               ; delta  = lambda * (H + lambda) $

Define the regressor vector for the primary equation. User supplies the set of names
and the name of y in the regression command to follow.

(*)    NAMELIST ; W = lambda,rhs in regression .... $

The next command gets the right coefficients, but the wrong standard errors.

       REGRESS  ; lhs = y ; rhs = W $

We now obtain the standard errors. First, we need estimates of rhoj and sigmaj. We
use the correction to the OLS estimate given by Heckman.

       CALC     ; thetajsq   = b(1) ^ 2
                ; sigmajsq   = sumsqdev / n  + thetajsq * xbr(delta)
                ; rhojsq     = thetajsq / sigmajsq $
```

```
The q variables in the next create statement must be set up by the user. If
you have selected on z=0, then qj = -p0 * pj for the J outcomes. If you have select-
ed on some other value, then for the J-1 outcomes not including 0 and not equal to
the selected one, qk = -pj * pk.  For the j'th one, qj = pj - pj^2. Then, each one
must be multiplied by delta when we compute F. Example: suppose there are four
choices, j=0,1,2,3. We select on j=2. Then, q1 = -delta*p1*p2, q2 = delta*p2*(1 -
p2), q3 = -delta*p2*p3. User supplies the format for these. Variable t is used to
compute W'(I-rhosq D)W simply as a weighted sum of squares and cross products.

       CREATE         ; t = 1 - rhojsq * delta
(*)                   ; q1 = ... ; q2 = ... ... $

The next set of matrix commands computes the corrected asymptotic covariance matrix
then displays the results using STAT.

(*)    MATRIX         ; F1   = xdot(W,V,q1) ; F2=xdot(W,V,q2) ; ...
                      ; F    = F1,F2,...
                      ; PSI  = sigmajsq * xdot(W,W,t)
                               & thetajsq * F | SG | F'
                      ; C    = xpxi(W) | PSI | xpxi(W)
                      ; stat(B,C) $
```

5 Semiparametrische selektionskorrigierte Schätzung

5.1 Modellspezifikation, semiparametrische Schätzung

Der zweistufige vorne vorgestellte Heckman-Ansatz ist ein parametrischer Ansatz, der für die Selektionsgleichung (PROBIT oder LOGIT/MNLOGIT) und für die Outputgleichung eine bestimmte Verteilungsfom der Fehlerterme (z.B. Normalverteilung) annimmt.

Ein vollständig nichtparametrischer Ansatz, ein Ansatz also, der keine explizite Verteilungsannahme vorsieht, und die Verteilungs- respektive Dichtefunktion bspw. über eine Kernel-Schätzung bestimmt, wird von Greene 2003, 704 als ‚fairly exotic' mit geringer empirischer Relevanz beschrieben. Hauptgrund ist die sehr beschränkte Berücksichtigung erklärender Größen.

Semiparametrische Ansätze ermitteln meist die Fehlerverteilung der Selektionsgleichung aus den Daten, behalten aber beispielsweise die Normalverteilungsannahme für die Outcomegleichung bei.

Ein Beispiel ist das Single Index Modell von Klein und Spady 1993, das von einer linearen ‚Regressions'-Formulierung ausgeht:

$$E\left[y_i \middle| x_i\right] = E\left[y_i \middle| \beta' x_i\right]$$

mit der Selektionsgleichung

$$\Pr ob(y_i = 1 | x_i) = F(\beta' x_i \mid x_i) = G(\beta' x_i)$$

wobei G(.) eine unbekannte Verteilungsfunktion ist, deren Gestalt – z.B. mittels einer nichtparametrischen Kernel-Schätzung - sowie die Parameter β zu schätzen sind.

Generell ist festzuhalten, dass sich die semiparametrischen Ansätze trotz ihrer Flexibilität hinsichtlich der Datenstruktur (und mehr noch die nicht parametrischen Ansätze) in empirischen Analysen nicht durchgesetzt haben, da erklärende Variablen nur sehr beschränkt – wenn überhaupt – eingesetzt werden können. In der Outcomegleichung führt zudem die semiparametrische Selektionsschätzung zu größeren Varianzen gegenüber den parametrischen Ansätzen (Greene 2003, S. 704ff).

5.2 Beispiel: A Comparison of Binary Choice Estimators

Gerfin (1996) did an extensive analysis of several binary choice estimators, the probit model, Klein and Spady's single index model, and Horowitz's smoothed maximum score estimator. (A fourth "seminonparametric" estimator was also examined, but in the interest of brevity, we confine our attention to the three more widely used procedures.) The several models were all fit to two data sets on labor force participation of married women, one from Switzerland and one from Germany. Variables included in the equation were (our notation), x_1 = a constant, x_2 = age, x_3 = age^2, x_4 = education, x_5 = number of young children, x_6 = number of older children, x_7 = log of yearly nonlabor income, and x_8 = a dummy variable for permanent foreign resident (Swiss data only). Coefficient estimates for the models are not directly comparable. We suggested in Example 19.1 that they could be made comparable by transforming them to marginal effects. Neither MSCORE nor the single index model, however, produces a marginal effect (which does suggest a question of interpretation). The author obtained comparability by dividing all coefficients by the absolute value of the coefficient on x_7. The set of normalized coefficients estimated for the Swiss data appears in Table 19.5, with estimated standard errors (from Gerfin's Table III) shown in parentheses.

Given the very large differences in the models, the agreement of the estimates is impressive. [A similar comparison of the same estimators with comparable concordance may be found in Horowitz (1993, p. 56).] In every case, the standard error of the probit estimator is smaller than that of the others. It is tempting to conclude that it is a more efficient estimator, but that is true only if the normal distribution assumed for the model is correct. In any event, the smaller standard error is the payoff to the sharper specification of the distribution. This payoff could be viewed in much the same way that parametric restrictions in the classical regression make the asymptotic covariance matrix of the restricted least squares estimator smaller than its unrestricted counterpart, even if the restrictions are incorrect.

Gerfin then produced plots of $F(z)$ for z in the range of the sample values of $\mathbf{b}'\mathbf{x}$. Once again, the functions are surprisingly close. In the German data, however, the Klein–Spady estimator is nonmonotonic over a sizable range, which would cause some difficult problems of interpretation. The maximum score estimator does not produce an estimate of the probability, so it is excluded from this comparison. Another comparison is based on the predictions of the observed response. Two approaches are tried, first counting the number of cases in which the predicted probability exceeds 0.5. ($\mathbf{b}'\mathbf{x} > 0$ for MSCORE) and second by summing the sample values of $F(\mathbf{b}'\mathbf{x})$.

TABLE 19.5 Estimated Parameters for Semiparametric Models

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	h
Probit	5.62	3.11	0.44	0.03	−1.07	−0.22	−1.00	1.07	—
	(1.35)	(0.77)	(0.10)	(0.03)	(0.26)	(0.09)	—	(0.29)	
Single index	—	2.98	−0.44	0.02	−1.32	−0.25	−1.00	1.06	0.40
	—	(0.90)	(0.12)	(0.03)	(0.33)	(0.11)	—	(0.32)	
MSCORE	5.83	2.84	−0.40	0.03	−0.80	−0.16	−1.00	0.91	0.70
	(1.78)	(0.98)	(0.13)	(0.05)	(0.43)	(0.20)	—	(0.57)	

(Once again, MSCORE is excluded.) By the second approach, the estimators are almost indistinguishable, but the results for the first differ widely. Of 401 ones (out of 873 observations), the counts of predicted ones are 389 for probit, 382 for Klein/Spady, and 355 for MSCORE. (The results do not indicate how many of these counts are correct predictions.)

Quelle: Greene 2004

6 Fractional Regression Models

Für viele ökonomische Bereiche ist die interessierende Variable oft ein Anteil, der nur im Intervall $[0 \leq y \leq 1]$ definiert ist. Beispiele sind Marktanteile von Firmen, Portfolioanteile, Ausgabenanteile etc. Diese Intervallbeschränkung, und auch das spezielle Interess für die Intervallgrenzen (zB. Ausgabenanteil von Null bzw. das ganze Einkommen wird ausgegeben) erfordert spezielle Schätzer. Das klassische lineare Rgressionsmodell ist dafür nicht geeignet, da es keine Garantie für vorhergesagte Werte innerhalb des Intervalls liefert.

Die Studie von Papke and Wooldridge (1996) gilt als 'seminal paper'. Panelmodelle für fractional regresseion models werden von Papke und Wooldridge (2008) diskutiert. Einen neueren Überblick über alternative Schätzstrategien geben Ramalho, Ramalho und Murteira 2011.

Die Situation für den angewandten Forscher haben Ramalho, Ramalho und Murteira (2011, p. 20) wie folgt beschrieben:
"Frequently, in applied work, researchers' main interest lies in the estimation of the conditional mean of y, given a set of regressors. In this case, practitioners face two main decisions:

1. which functional form to assume for the conditional expectation of y and
2. which method to employ in the estimation of the resulting model.

In addition, in the case of boundary observations, practitioners have also to decide whether one- or two-part models should be used. So far, most authors have chosen to assume a logistic form for the conditional mean of y, without assessing whether alternative functional forms are more appropriate, and to use the robust quasi-maximum likelihood (QML) method suggested by Papke and Wooldridge (1996), without checking whether a more efficient estimation method could be used. However, in both cases, there are a number of alternatives that may be employed and various simple test procedures that may be used to assess their adequacy. Similarly, the option between a single- and a two-part model is usually made apriori and, as far as we know, has never been tested.
In this paper we survey the main alternative regression models and estimation methods that are available for dealing with fractional response variables and propose a full testing methodology to assess the validity of the assumptions required by each estimator."

V PANELMODELLE

1 Modellstruktur der 'discrete choice'-Panelmodelle

Paneldaten werden aus wiederholten Beobachtungen der gleichen Einheiten über die Zeit gewonnen. In den Wirtschafts- und Sozialwissenschaften sind das in der Regel wiederholte Befragungen der gleichen Mikroeinheiten. In diesen Daten wird somit ein relativ großer Querschnitt mit den gleichen Mikroeinheiten (Personen, Firmen) über mehrere Perioden betrachtet. Beispiele sind das Sozio-ökonomische Panel (SOEP), begonnen durch den Sonderforschungsbereich 3 ‚Mikroanalytische Grundlagen der Gesellschaftspolitik' der Universitäten Frankfurt und Mannheim und nun vom DIW in Berlin weitergeführt, mit nun ca. 8.500 Haushalten und über 20.000 Personen, das Hannoveraner Firmenpanel oder die Michigan Panel Study of Income Dynamics (PSID), das mit ca. 6.000 Familien und ca. 15.000 Personen bereits 1968 begonnen wurde.

Mit Schwerpunkt auf vielen Querschnittsbeobachtungen in den Wirtschafts- und Sozialwissenschaften spielt die Heterogenität über die Mikroeinheiten eine besondere Rolle.

Für die explizite Berücksichtigung der dieser Heterogenität sowie des zeitlichen Einflusses sind generell fixed und random effects Modelle entwickelt worden. Greene (2000, chapt. 14) bspw. behandelt allgemein den Fall linearer Panelmodelle.

Im Gegensatz zu den linearen Panelmodellen ist bei den 'discrete choice'-Panelmodellen als nichtlineare Modelle eben die besondere Form der linken Seite der Regressionsgleichung zu berücksichtigen.

Literatur zur Panelanalyse: Baltagi 2005, Hsiao 2003, Wooldridge 2002, Arellano 2002, Chay und Hyslop 2014.

2 Fixed effects 'discrete choice' LOGIT und PROBIT Panelmodelle

Bei den fixed effects Panelmodellen wird davon ausgegangen, dass die Differenzen zwischen den Mikroeinheiten mit Differenzen im konstanten Term eingefangen werden können:

$$y_i = i\alpha_i + \beta' x_i + \varepsilon_i$$

wobei y_i und x_i die T Beobachtungen der i-ten Mikroeinheit beinhalten.

Ein **fixed effects LOGIT Panelmodell**, das die Heterogenität über die Mikroeinheiten berücksichtigt ist

$$P[y_{it}=1] = \frac{e^{\alpha_i+\beta' x_i}}{1+e^{\alpha_i+\beta' x_i}} = \frac{\exp(\alpha_i+\beta' x_i)}{1+\exp(\alpha_i+\beta' x_i)}$$

In diesem nichtlinearen Modell ist es nun nicht mehr wie im linearen Modell möglich, die Heterogenität über Differenzen oder Abweichungen vom Gruppenmittel, herauszurechnen.

Einen Schätzansatz dafür hat Chamberlain 1980 mit der Maximierung einer konditionalen Likelihood Funktion entwickelt. Ob die Heterogenität signifikant ist, kann dann mit einem Hausman Test über die Differenz zwischen unkonditionierten und konditionierten Schätzansatz getestet werden (Greene 2000, S. 841).

Für das **fixed effects PROBIT Modell** mit

$$y_{it}^{*} = \beta' x_{it} + \varepsilon_{it}, \qquad \varepsilon_{it} \sim N(0,1), i = 1,...,n, t = 1,...,T_i$$

$$y_{it} = 1 \quad wenn \; y_{it}^{*} > 0, und \; 0 \; sonst$$

gibt es im Gegensatz zum fixed effects LOGIT Modell keine geeignete Lösung, um die α_i zu schätzen.

3 Random effects 'discrete choice' LOGIT und PROBIT Panelmodelle

Die random effects Modelle zeichnen sich dadurch aus, dass die individuenspezifischen konstanten Terme zufällig (randomly) über die Mikroeinheiten verteilt sind

$$y_{it} = \alpha + \beta' x_{it} + u_i + \varepsilon_{it}.$$

Idealerweise sollten die ε_{it} und ε_{is} frei korreliert sein innerhalb, aber nicht über die Gruppen hinweg. Das führt aber zu Rechenproblemen bezüglich gemeinsamer Verteilungen einer T_i-variaten Normalverteilung (PROBIT-Ansatz). Eine etwas weniger allgemeine Spezifikation haben Butler und Moffitt (1982) für random effects Modelle mit diskreten und beschränkt abhängigen Variablen entwickelt.

4 Fixed effects vs. Random effects LOGIT und PROBIT Panelmodelle

Maddala 1987 ist der Frage nachgegangen, welche LOGIT und/oder PROBIT Panelmodelle unter welchen Bedingungen heranzuziehen sind. Er ist zu folgenden Ergebnissen gekommen (Maddala 1987, S: 314-315):

(i) Fixed effects logit and probit models. Here we argue that for large N and small T (which is typically the case) ML estimation of the fixed effects model gives inconsistent estimates of the parameters. However, in the logit model one can obtain consistent estimates using the conditional ML method (conditioning on the fixed effects). Such conditioning is not possible with the probit model. Hence, for the analysis of the fixed effects model, the logit model is the appropriate one.

(ii) Random effects logit and probit models. For the analysis of random effects, the probit model is the appropriate one. The random effects produce correlations among the errors and the multivariate logistic distribution is too restrictive for this purpose (since it implies that all correlations are 0.5). The probit model is based on the multivariate normal distribution which is more flexible. For the analysis of random effects probit models we consider three models suggested by Heckman and Willis (1976), Avery, Hansen, and Hotz (1983), and Chamberlain (1985).

(iii) Fixed effects and random effects tobit models. These are extensions of the previous models. Illustrative examples are in Heckman and MaCurdy (1980) and Hausman and Wise (1979).

(iv) Autoregressive logit and probit models. These are extensions of models in (i) and (ii) to the case where lagged y's are included as explanatory variables. The conditional logit model can be extended to this case but only under some restrictive assumptions. The random effects probit model can be easily extended to this case.

Weitere Vertiefung: Greene 2003, chapt. 21.5 inklusive semiparametrischer und nichtparametrischer Schätzansätze, Cameron und Trivedi 2005 chapt. 23 sowie Moosmüller 2004, Kapitel 2.4.5 und Chay und Hyslop 2014.

PROGRAMMPAKETE

ET: Econometrics Toolkit, Greene, W.H. (1990), ET - The Econometrics Toolkit, Version 2, User's Guide, Econometric Software, Inc., New York

LIMDEP: Limited dependent variables, Greene, W.H. (1990, 1992, 1995, 2004), LIMDEP User's Guide (Version 5.1, 6.0, 7.0, 8.0), Econometric Software Inc., New York, www.limdep.com

ET, LIMDEP: William H. Greene, Stern Graduate School of Business, New York University, 100 Trinity Place, New York, NY 1006

EALIMDEP: LIMDEP for students (download von http:/ffb.uni-lueneburg.de)

STATA: Allgemeines Statistisches Programmpaket, Version 8, Stata Corporation, www.stata.com

EViews: Econometric package, Quantitative Micro Software (1997), EViews User's Guide, Irvine, California, USA, www.eviews.com

GAUSS: Statistisches Paket, prozedural, Metasprache Aptech Systems Inc. 26250 196th Place SE, Kent, WA 98042, (206)-631-667), www.aptech.com

GLAMOUR: Fahrmeir, L., Frost, H., Hennevogel, W., Kaufmann, H., Kranert, T. und G. Tutz (1990), User's Guide for GLAMOUR, Version 2.0, Institut für Statistik und Wirt schaftsgeschichte, Universität Regensburg und München

LITERATUR

Agresti, A. (2007), An Introduction to Categorical Data Analysis, 2nd edition, John Wiley, New Jersey

Alba, R. (1988), Interpreting the Parameters of Log-Linear Models, in: J.S. Long (ed.) (1988), Common Problems Proper Solutions – Avoiding Error in Quantitative Research, Sage Publications,Newbury Park, 258-287

Allenby, G. und J. Ginter (1995), The Effects of InStore Displays and Feature Advertising on Consideration Sets, in: International Journal of Research in Marketing, 12, 67-80

Amemiya T. (1985), Advanced Econometrics, Basil Blackwell, Oxford

Amemiya, T. (1975), Qualitative Response Models, Annals of Economic and Social Measurement

Amemiya, T. (1981), Qualitative Response Models: A Survey, Journal of Economic Literature 19, 483-536

Anderson D.R., Sweeney D.J., T.A. Williams, J. Freeman und E. Shoesmith (2010), Statistics for Business and Economics, Minnesota.

Andreß, H.-J., Hagenaars, J.A. und S. Kühnel (1997), Analyse von Tabellen und kategorialen Daten, Springer-Verlag, Berlin, Heidelberg.

Arellano, M. (2003), Panel Data Econometrics, Oxford University Press, Oxford.

Arminger, G. und U. Küsters (1986), Statistische Verfahren zur Analyse qualitativer Variablen, Bergische Universität Wuppertal, Bergisch Gladbach

Auer, L.v. (2011), Ökonometrie, Eine Einführung, 5. Auflage, Springer Verlag, Berlin.

Baltagi, B.H. (2008), Econometric Analysis of Panel Data, 4th Edition, John Wiley & Sons, Chichester.

Ben-Akiva, M. und S.R. Lerman (1985), Discrete Choice Analysis, MIT Press, Cambridge (Mass.)

Berkson, J. (1944), Application of the Logistic Function to Bio-Assay, in: Journal of the American Statistical Association 39, 357-365

Bhat, C. (1945), A Heteroskedastic Extreme Value Model of Intercity Mode Choice, in: Transportation Research, 30/1, 16-29

Blossfeld, H.-P., Hamerle, A. und K.U. Mayer (1989), Ereignisanalyse - Statistische Theorie und Anwendung in den Wirtschafts- und Sozialwissenschaften

Blundell, R. (o.J.), Microeconometrics Lewcture Notes - 1: Some Notes on Discrete Choice and Limited Dependent Variable Models, University College London, London

Börsch-Supan, A. (1987), Econometric Analysis of Discrete Choice. With Applications on the Demand for Housing in the U.S. and West-Germany, Springer, Heidelberg/New York

Brant, R. (1990), Assessing proportional odds model for ordinal logistic regression, in: Biometrica, Vol. 46, 1171-1178.

Breen, R. (1996), Regression Models – Censored, Sample-Selected or Truncated Data, Sage University Press, Thousand Oaks, New Delhi.

Brown, R.L., Durbin, J. and J.M. Evans (1975), Techniques for Testing the Constancy of Regression Relationships over Time, RSS, Vol. 37, 149-192

Butler, J. und R. Moffitt (1982), A Computationally Efficient Quadrature Procedure for the One Factor Multinomial Probit Model, in: Econometrica 50, 761-764

Cameron, A. C. and P. K. Trivedi (2010), Microeconometrics Using Stata, Revised Edition, Stata Press

Cameron, A. C. and P. K. Trivedi (2005), Microeconometrics – Methods and Applications, Cambridge University Press, New York

Chamberlain, G. (1980), Analysis of Covariance with Qualitative Data, in: Review of Economic Studies 47,

Chamberlain, G. (1985), Panel Data, Griliches, Z. und M.D. Intriligator (Hg.), Handbook of Econometrics, Vol. 2, Chapt. 24, 1247-1320, North Holland, Amsterdam

Chay, Kenneth, Y. and Dean R. Hyslop (2014), Identification and Estimation of Dynamic Binary Response Panel Data Models: Empirical Evidence Using Alternative Approaches, in: Safety Nets and Benefit Dependence, Research in Labor Economics, Volume 39, 1-39, http://dx.doi.org/10.1108/S0147-912120140000039001

Chow, G.C. (1983), Econometrics (Chapt. 8), McGraw-Hill, Auckland/Tokyo

Collett, D. (2003), Modelling Binary Data, 2nd ed., Chapman & Hall/CRC, Boca Raton.

Daganzo, C. (1980), Multinomial Probit, New York

Davidson, R. and J. G. MacKinnon (2004), Econometric Theory and Methods, Oxford University Press, New York

Deaton, A. und J. Muellbauer (1980), Economics and Consumer Behavior, Cambridge, Sidney

Desu, M.M. und D. Raghavarao (2004), Nonparametric Statistical Methods for Complete and Censored Data, Chapman & Hall/CRC, Boca Raton.

Dhrymes, Ph. J. (1986), Limited Dependent Variables, Griliches, Z. und M.D. Intriligator (Hg.), Handbook of Econometrics, Vol. 3, Chapt. 27, 1567-1632, North Holland, Amsterdam

Eckey, H.-Fr., Kosfeld, R. und Chr. Dreger (2001), Ökonometrie, Grundlagen – Methoden – Beispiele, 2. Aufl., Wiesbaden.

Fahrmeir, L., Kneib, Th. und S. Lang (2007, 2009), Regression, Modelle, Methoden und Anwendungen, Berlin.

Fair, R. (1978), A Theory of Extramarital Affairs, in Journal of Political Economy, 86, 45-61

Goodman, Leo A. (1984), The Analysis of Cross-Classified Data Having Ordered Categories, Cambridge (Mass.), Harvard University Press

Gourieroux, Chr. (2000), Econometrics of Qualitative Dependent Variables, Cambridge, Cambridge University Press.

Greene W.H. and D.A. Hensher (2010), Modeling ordered choices – A primer, Cambridge University Press, Cambridge.

Greene, W.H. (2004), LIMDEP Programmpaket, Version 8, New York University

Greene, W.H. (2008a, 2003, 2000, 1997, 1993, 1990), Econometric Analysis, Editions 1-6, New York/London.

Greene, W.H. (2008b) Modeling Ordered Choices: A Primer and Recent Developments, http://ssrn.com/abstract=1213093.

Greene, W.H. and D.A. Hensher (2010), Modeling ordered choices – A Primer, Cambridge University Press, Cambridge.

Hagenaars, J.A. (1990), Categorical Longitudinal Data – Log-Linear Panel, Trend, and Cohort Analyses, Sage Publications, Newburry Park/London/New Delhi.

Hangelberger, D. (2011), Arbeitszufriedenheit und Flexibilität – Internationaler Vergleich und Adaptions- und Antizipationseffekte, Dissertation, Lüneburg.

Hausman, J. and D. McFadden (1984), A Specification Test for the Multinomial Logit Model, in: Econometrica, 52, 1219-1240.

Heckman, J.J. (1976), The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models, in: Annals of Economic and Social Measurement, 5, 475-492

Heckman, J.J. (1979), Sample Selection Bias as a Specification Error, Econometrica, 153-161

Heckman, J.J. und T.E. MaCurdy (1986), Labor Econometrics, Griliches, Z. und M.D. Intriligator (Hg.), Handbook of Econometrics, Vol. 3, Chapt. 32, North Holland, Amsterdam

Hosmer D.W. and S. Lemeshow (2000), Applied Logistic Regression, 2nd Edition, John Wiley & Sons, New York.

Hsiao, Ch. (2003), Analysis of Panel Data, 2nd ed., Cambridge University Press, Cambridge

Hsiao, Ch., Lahiri, K., Lee, L.-F. und M.H. Pesaran (1999), Analysis of panels and limited dependent variable models, Cambridge University Press, Cambridge.

Hübler, O. (1989), Ökonometrie, UTB-Taschenbücher 1538, Stuttgart/New York

Hübler, O. (2005), Einführung in die empirische Wirtschaftsforschung, R.Oldenbourg Verlag, München.

Hübler, O. (2007), Übungsbuch zur empirischen Wirtschaftsforschung, Aufgaben, Wiederholungsfragen, Ergänzungen und Lösungen, R.Oldenbourg Verlag, München.

Joesch, J.M. and K. Spiess (2006), European mothers' time spent looking after children - Differences and similarities across nine countries, in: electronic International Journal of Time Use Research, 3/1, 1-27

Judge, G.G., Griffiths, W.E., Carter Hill, R., Lütkepohl, H. und T.-C. Lee (1985), The Theory and Practice of Econometrics, 2nd ed., New York/Singapore

Kasprzyk, D., Duncan, G., Kalton, Gr. Und M.P. Singh, John Wiley & Sons, New York, Chichester, Brisbane, Toronto, Singapore

Lechner, M. (1995), Ökonometrische Modelle für qualitative Paneldaten, Das Beispiel der selbständigen Erwerbstätigkeit in den neuen Bundesländern, Campus Verlag, Frankfurt.

Lee, M.-J. (2005), Micro-econometrics for Policy, Program, and Treatment Effects, Oxford University Press: Oxford

Lluch, C., The Extended Linear Expenditure System, in: European Economic Review 4, 21-32

Long, J.S. (1997), Regression Models for Categorical and Limited Dependent Variables, Sage Publications, Thousand Oaks/London/New Delhi

Long, J.S. (1997), Regression models for categorical and limited dependent variables, Sage Publications, Thousand Oaks, California.

Long, J.S. (Ed.) (1988), Common Problems/Proper Solutions – Avoiding Error in Quantitative Research, Sage Publications, Newbury Park/New Delhi

Long, J.S. und J. Freese (2006), Regression Models for Categorical Dependent Variables Using Stata, 2nd ed., Stata Press, Texas

Long, J.S. und J. Freese (2006), Regression models for categorical dependent variables using Stata, 2. Auflage, Stata Press, College Station, Texas.

Maddala, G.S. (1983), Limited-dependent and qualitative variables in econometrics, Cambridge University Press, Cambridge

Maddala, G.S. (1987), Limited Dependent Variable Models Using Panel Data, Jorurnal of Human Resources 22, 307-338

Maddala, G.S. (2001), Intoduction to Econometrics, 3. Auflage, Macmillan/New York

Maier, G. und P. Weiss (1990), Modelle diskreter Entscheidungen - Theorie und Anwendung in den Sozial- und Wirtschaftswissenschaften, Springer Verlag, Wien/New York

Mansky, C.F. und D. McFadden (Hg.) (1981), Structural Analysis of Discrete Data with Econometric Applications, MIT Press, Cambridge (Mass.)

Marcus, A. and W. Greene (1985), The Determinants of Rating Assignment and Performance, Working Paper CRC528 Center for Naval Analyses

McFadden, D. (1974), Conditional Logit Analysis of Qualitative Choice Behaviour, Zarembka, P. (Hg.), Frontiers in Econometrics

McFadden, D. (1976), Quantal Choice Analysis: A Survey. Annals of Economic and Social Measurement, 5, 363-86.

McFadden, D. (1985), Econometric Analysis of Qualitative Response Models, Griliches, Z. und M.D. Intriligator (Hg.), Handbook of Econometrics, Vol. 2, Chapt. 24, 1395-1459, North Holland, Amsterdam

McFadden, D. (1989), A Method of Simulated Moments for Estimation of Multinomial Probits without Numerical Integration, Econometrica 57, 995-1026

Merz, J. (1983), Der Einfluß sozioökonomischer Größen auf die individuelle private Nachfrage nach dauerhaften Konsumgütern - Eine Anwendung der diskreten Entscheidungsmodelle LOGIT und TOBIT, in: Zeitschrift für Wirtschafts- und Sozialwissenschaften, 103. Jg., Heft 3, S. 225-253

Merz, J. (1983), FELES: The Functionalized Extended linear Expenditure System – Theory, Estimation Procedures and Application to Individual Household Consumption Expenditures Involving Socioeconomic and Sociodemographic Characteristics., in: European Economic Review, No. 23 (1983), S. 359-394

Merz, J. (1987), Das Arbeitsangebot verheirateter Frauen in der Bundesrepublik Deutschland - Eine mikroökonometrische Analyse unter Berücksichtigung des 'selectivity bias', in: Krupp, H.J., Hanefeld, U. (Hg.), Lebenslagen im Wandel, Frankfurt/New York, S. 126-157

Merz, J. (1989), Markt- und nichtmarktmäßige Aktivitäten privater Haushalte - Theoretischer Ansatz, repräsentative Mikrodaten, mikroökonometrische Analyse und Mikrosimulation wirtschafts- und sozialpolitischer Maßnahmen für die Bundesrepublik Deutschland (Habilitationsschrift), Frankfurt

Merz, J. (1989), Zum Erwerbsverhalten von Frauen im Haushaltszusammenhang - Eine Mikroanalyse von Lohnsatz- und Arbeitszeitdeterminanten mit Daten des Sozio-ökonomischen Panels, in: Emmerich, K., Hardes, H.-D., Sadowski, D. und E. Spitznagel (Hg.), Einzel- und gesamtwirtschaftliche Aspekte des Lohnes, Nürnberg, S. 133-154

Merz, J. (1990), Female Labor Supply in the Household Context - Estimates of Wage Rate and Working Hours Based on a Two Stage TOBIT Approach with Data from the German Socio-economic Panel, in: Applied Econometrics Association (Hg.), Proceedings of the 31st International Conference of the Applied Econometrics Association, 5.-7.Dezember 1990, Straßbourg (Frankreich), Brüssel (Belgien), 1991a, S. 62-82

Merz, J. (1990), Female Labor Supply: Labor Force Participation, Market Wage Rate and Working Hours of Married and Unmarried Women in the Federal Republic of Germany - The Influence of Occupational Career, Transfers and Socio-Economic Variables Based on the Socio-Economic Panel, in: Jahrbücher für Nationalökonomie und Statistik, S. 240-270

Merz, J. (1990), Time Allocation in Market and Non-Market Activities - Generalized Discrete/Continuous MNL/COLS Estimates of Multiple Labor Supply, Sfb 3-Working Paper No. 337, Frankfurt/Mannheim

Merz, J. (1992), Zur Dynamik markt- und nichtmarktmäßigen Arbeitsangebots: Zeitverwendung verheirateter Frauen in Beruf und Haushalt - Eine Panelanalyse, in: Hujer, R., Schneider, H. und W. Zapf (Hg.), 1992, Herausforderungen an den Wohlfahrtsstaat im strukturellen Wandel, Campus Verlag, Reihe Wirtschaftswissenschaften, Bd. 23, Frankfurt/New York, S. 343-378

Merz, J. (1996), Market and Non-Market Labour Supply and the Impact of the Recent German Tax Reform - Incorporating Behavioural Response, in: Ann Harding (Hg.) Microsimulation and Public Policy, (IARIW Special Conference on Microsimulation and Public Policy, Canberra, Australia, 5.-9.12.93), Amsterdam, S. 177-202

Merz, J. (2012), Markt- und nichtmarktmäßige Aktivitäten privater Haushalte - Mikrotheorie, Mikrodaten, Mikroökonometrie und Mikrosimulation für die Bundesrepublik Deutschland, 2. Auflage, LIT-Verlag. Berlin

Merz, J. and T. Rathjen (2012), Intensity of Time and Income Interdependent Multidimensional Poverty: Well-Being and Minimum 2DGAP – German Evidence, Forschungsinstitut Freie Berufe (FFB), Leuphana Universität Lüneburg, Revised FFB-Discussion Paper No.92 (2011), Lüneburg.

Merz, J. und Klaus G. Wolff (1993), The Shadow Economy: Illicit Work and Household Production - A Microanalysis of West Germany, in: The Review of Income and Wealth, Vol. 39, No. 2, S. 177-194

Merz, J. und P. Paic (2005), Zum Einkommen der Freien Berufe – Eine Ordered Probit-Analyse ihrer Determinanten auf Basis der FFB-Onlineumfrage, in: Schulte, R. (Hg.), Ergebnisse der MittelstandsForschung, Lit-Verlag, Münster/London, 327-348

Merz, J., Hanglberger, D. und R. Rucha (2009), The Timing of Daily Demand for Goods and Services – Multivariate Probit Estimates and Microsimulation Results for an Aged Population with German Time Use Diary Data, FFB Discussionpaper No. 77, Lüneburg

Merz, J., Hanglberger, D. und R. Rucha (2010), The Timing of Daily Demand for Goods and Services – Microsimulation Policy Results of an Aging Society, Increasing Labour Market Flexibility, and Extended Public Childcare in Germany, in: Journal of Consumer Policy, Special Issue: Time Allocation, Consumption, and Consumer Policy, 33, 119-141, 2010

Mittelhammer; R. C., Judge G.J. and J.M. Douglas (2000), Econometric Foundations, Cambridge.

Moosmüller, G. (2004), Merthoden der empirischen Wirtschaftsforschung, München

Papke, L.E. and J.M. Wooldridge (1996), Econometric methods for fractional response variables with an application 401(k) plan participation rates, in: Journal of Applied Econometrics, Vol. 11, 6190-632.

Papke, L.E. and J.M. Wooldridge (2008), Panel data methods for fractional response variables with an application to test pass rates, in Journal of Econometrics, Vol. 45, 121-133.

Peracchi, Fr. (2001), Econometrics, John Wiley & Sons, Chichester/Toronto

Pindyck, R.S. und D. L. Rubinfeld (1997), Econometric and Models - Economic Forecasts, 4th ed., New York/Toronto

Powers, D.A. und Yu Xie (2000), Statistical Methods for Categorical Data Analysis, Academic Press, London/San Diego

Pudney, S. (1989), Modelling Individual Coice - The Econometrics of Corners, Kinks and Holes, Basil Blackwell, Oxford

Pudney, St. (1989), Modelling Individual Choice – The Econometrics of Corners, Kinks and Holes, Basil Blackwell, Oxford, New York

Quester, A. and W. Greene (1982), Divorce Risk and Wives' Labor Supply Behavior, in: Social Science Quarterly, 63, 16-27

Ramalho, E.A., Ramalho, J.J.S. and J.M.R. Murteira (2011), Alternative Estimating and Testing Empirical Strategies for Fractional Response Models, in: Journal of Economic Surveys, Vol. 25, 19-68, doi: 10.1111/j.1467-6419.2009.00602.x.

Ronning, G. (1991), Mikroökonometrie, Springer Verlag, Berlin/Barcelona

Schmidt, P. und R. Strauss (1975a), The Prediction of Occupation Using Multinomial Logit Models, in: International Economic Review, 16, 471-486

Schmidt, P. und R. Strauss (1975b), Estimation of Models with Jointly Dependent Qualitative Variables: A Simultaneous Logit Approach, in: Econometrica, 43, 745-755

Sigelman, L. und L. Zeng (1999), Analyzing Censored and Sample-Selected Data with Tobit and Heckit Models, in: Political Analysis, 8/2, 167-182

Stewart, J. (2009), Tobit or Not Tobit?, IZA Discussion Paper No. 4588, Bonn.

Stone, R. (1954), Linear Expenditure Systems and Demand Analysis – An Application to the Patterns of British Demand, in: Economic Journal 64, 511-527

Tobin, J. (1958), Estimation of Relationships for Limited Dependent Variables, Econometrica, 24-36

Tutz, G. (2000), Die Analyse kategorialer Daten, Anwendungsorientierte Einführung in Logit-Modellierung und kategoriale Regression, R. Oldenbourg Verlang, München.

Urban, D. (1993) Logit Analyse, Statistische Verfahren zur Analyse von Modellen mit qualitativen Response-Variablen, Gustav Fischer Verlag, Stuttgart.

Veall, M.R. und K.F. Zimmermann (1994), Goodness of FitMeasures in the Tobit Model, in: Oxford Bulletin of Economics and Statistics, 56/4, 485-499

Verbeek, M. (2008), A Guide to Modern Econometrics, 3rd ed,, John Wiley & Sons, Chichester.

Vogelvang, B. (2005), Econometrics – Theory and Applications with EViews, Pearson/Addison Wesley, Harlow.

Winkelmann, R. (1997), Econometric Analysis of Count Data, Second, Revised and Enlarged Edition, Springer-Verlag, Berlin.

Winkelmann, R. und St. Boes (2009), Analysis of Microdata, 2nd edition, Springer-Verlag, Berlin Heidelberg.

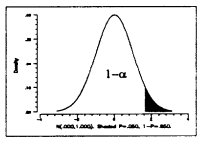
Wooldridge, J.M. (2002), Econometric Analysis of Cross Section and Panel Data, The MIT Press, Cambridge, Mass.

Wooldridge, J.M. (2006), Introductory Econometrics: A Modern Approach, Third edition, Thomson South-Western.

Zellner, A. und T. Lee (1965), Joint Estimation of Relationships Involving Discrete Random Variables, Econometrica 33, 383-394

VERTEILUNGSTABELLEN

Verteilungsfunktion $\Phi(z)$ der Standardnormalverteilung $N(0,1)$ (Quelle: Hartung et al. 1982, S. 734)



z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990

Ablesebeispiel: $\Phi(1,56) = 0,9406$; $\Phi(z) = 1-\alpha$

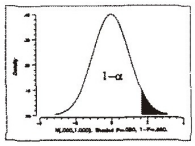
Erweiterung der Tafel: $\Phi(-z) = 1-\Phi(z)$

Approximation nach Hastings für $z > 0$

$$\Phi(z) \cong 1 - \frac{1}{\sqrt{2\cdot\pi}} \cdot e^{-\frac{z^2}{2}} \cdot \left(a_1 \cdot t + a_2 \cdot t^2 + a_3 \cdot t^3 + a_4 \cdot t^4 + a_5 \cdot t^5\right) \text{ mit } t = \frac{1}{1+b\cdot z},$$

$b = 0,2316419$, $a_1 = 0,31938153$, $a_2 = -0,356563782$, $a_3 = 1,781477937$,
$a_4 = -1,821255978$, $a_5 = 1,330274429$.

Quantile z der Standardnormalverteilung
$N(0,1)$ (Quelle: Hartung et al. 1982, S. 735)



$\Phi(z)$	z	$\Phi(z)$	z	$\Phi(z)$	z	$\Phi(z)$	z
0,9999	3,7190	0,9975	2,8071	0,965	1,8119	0,83	0,9542
0,9998	3,5401	0,9970	2,7478	0,960	1,7507	0,82	0,9154
0,9997	3,4316	0,9965	2,6968	0,955	1,6954	0,81	0,8779
0,9996	3,3528	0,9960	2,6521	0,950	1,6449	0,80	0,8416
0,9995	3,2905	0,9955	2,6121	0,945	1,5982	0,79	0,8064
0,9994	3,2389	0,9950	2,5758	0,940	1,5548	0,78	0,7722
0,9993	3,1947	0,9945	2,5427	0,935	1,5141	0,76	0,7063
0,9992	3,1559	0,9940	2,5121	0,930	1,4758	0,74	0,6433
0,9991	3,1214	0,9935	2,4838	0,925	1,4395	0,72	0,5828
0,9990	3,0902	0,9930	2,4573	0,920	1,4051	0,70	0,5244
0,9989	3,0618	0,9925	2,4324	0,915	1,3722	0,68	0,4677
0,9988	3,0357	0,9920	2,4089	0,910	1,3408	0,66	0,4125
0,9987	3,0115	0,9915	2,3867	0,905	1,3106	0,64	0,3585
0,9986	2,9889	0,9910	2,3656	0,900	1,2816	0,62	0,3055
0,9985	2,9677	0,9905	2,3455	0,890	1,2265	0,60	0,2533
0,9984	2,9478	0,9900	2,3263	0,880	1,1750	0,58	0,2019
0,9983	2,9290	0,9850	2,1701	0,870	1,1264	0,56	0,1510
0,9982	2,9112	0,9800	2,0537	0,860	1,0803	0,54	0,1004
0,9981	2,8943	0,9750	1,9600	0,850	1,0364	0,52	0,0502
0,9980	2,8782	0,9700	1,8808	0,840	0,9945	0,50	0,0000

Ablesebeispiel: $z_{0,95} = 1,6449$

Erweiterung der Tafel: $z_{1-\alpha} = -z_{\alpha}$

Approximation nach Hastings für $0,5 < \Phi(z) < 1$

$$z \cong t - \frac{a_0 + a_1 \cdot t + a_2 \cdot t^2}{1 + b_1 \cdot t + b_2 \cdot t^2 + b_3 \cdot t^3} \text{ mit } t = \sqrt{-2 \cdot \ln\left(1 - \Phi(z)\right)},$$

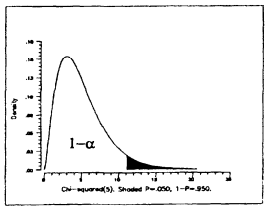
$a_0 = 2,515517,$ $a_1 = 0,802853,$ $a_2 = 0,010328,$
$b_1 = 1,432788,$ $b_2 = 0,189269,$ $b_3 = 0,001308.$

Häufig vorkommende z-Werte für die Bestimmung von Konfidenzintervallen

Z ist standardnormalverteilt		Konfidenzintervall	
		zweiseitig	einseitig
Häufig vorkommende z-Werte für:	$\alpha = 0,01$	$z = 2,58$	$z = 2,33$
	$\alpha = 0,05$	$z = 1,96$	$z = 1,65$
	$\alpha = 0,1$	$z = 1,65$	$z = 1,28$

χ^2-Verteilung für $P\left(O < \chi^2 \leq \chi^2_\alpha\right) = 1-\alpha$

(Quelle: J. Schwarze: Grundlagen der Statistik II, 1993, S. 306 - 307):

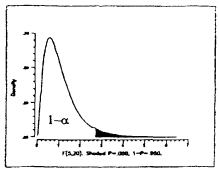


ν	1 − α										
	0,001	0,005	0,010	0,025	0,050	0,100	0,900	0,950	0,975	0,990	0,995	0,999
1	0,000	0,000	0,000	0,001	0,004	0,016	2,706	3,841	5,024	6,635	7,879	10,828
2	0,002	0,010	0,020	0,051	0,103	0,211	4,605	5,991	7,378	9,210	10,597	13,816
3	0,024	0,072	0,115	0,216	0,352	0,584	6,251	7,815	9,348	11,345	12,838	16,266
4	0,091	0,207	0,297	0,484	0,711	1,064	7,779	9,488	11,143	13,277	14,860	18,467
5	0,210	0,412	0,554	0,831	1,145	1,610	9,236	11,070	12,833	15,086	16,750	20,515
6	0,381	0,676	0,872	1,237	1,635	2,204	10,645	12,592	14,449	16,812	18,548	22,458
7	0,598	0,989	1,239	1,690	2,167	2,833	12,017	14,067	16,013	18,475	20,278	24,322
8	0,857	1,344	1,646	2,180	2,733	3,490	13,362	15,507	17,535	20,090	21,955	26,125
9	1,152	1,735	2,088	2,700	3,325	4,168	14,684	16,919	19,023	21,666	23,589	27,877
10	1,479	2,156	2,558	3,247	3,940	4,865	15,987	18,307	20,483	23,209	25,188	29,588
11	1,834	2,603	3,053	3,816	4,575	5,578	17,275	19,675	21,920	24,725	26,757	31,264
12	2,214	3,074	3,571	4,404	5,226	6,304	18,549	21,026	23,337	26,217	28,300	32,910
13	2,617	3,565	4,107	5,009	5,892	7,042	19,812	22,362	24,736	27,688	29,819	34,528
14	3,041	4,075	4,660	5,629	6,571	7,790	21,064	23,685	26,119	29,141	31,319	36,123
15	3,483	4,601	5,229	6,262	7,261	8,547	22,307	24,996	27,488	30,578	32,801	37,697
16	3,942	5,142	5,812	6,908	7,962	9,312	23,542	26,296	28,845	32,000	34,267	39,252
17	4,416	5,697	6,408	7,564	8,672	10,085	24,769	27,587	30,191	33,409	35,718	40,790
18	4,905	6,265	7,015	8,231	9,390	10,865	25,989	28,869	31,526	34,805	37,156	42,312
19	5,407	6,844	7,633	8,907	10,117	11,651	27,204	30,144	32,852	36,191	38,582	43,820
20	5,921	7,434	8,260	9,591	10,851	12,443	28,412	31,410	34,170	37,566	39,997	45,315
21	6,447	8,034	8,897	10,283	11,591	13,240	29,615	32,671	35,479	38,932	41,401	46,797
22	6,983	8,643	9,542	10,982	12,338	14,041	30,813	33,924	36,781	40,289	42,796	48,268
23	7,529	9,260	10,196	11,689	13,091	14,848	32,007	35,172	38,076	41,638	44,181	49,728
24	8,085	9,886	10,856	12,401	13,848	15,659	33,196	36,415	39,364	42,980	45,558	51,179
25	8,649	10,520	11,524	13,120	14,611	16,473	34,382	37,652	40,646	44,314	46,928	52,620

ν	$1-\alpha$											
	0,001	0,005	0,010	0,025	0,050	0,100	0,900	0,950	0,975	0,990	0,995	0,999
26	9,222	11,160	12,198	13,844	15,379	17,292	35,563	38,885	41,923	45,642	48,290	54,052
27	9,803	11,808	12,879	14,573	16,151	18,114	36,741	40,113	43,195	46,963	49,645	55,476
28	10,391	12,461	13,565	15,308	16,928	18,939	37,916	41,337	44,461	48,278	50,993	56,892
29	10,986	13,121	14,256	16,047	17,708	19,768	39,087	42,557	45,722	49,588	52,336	58,301
30	11,588	13,787	14,953	16,791	18,493	20,599	40,256	43,773	46,979	50,892	53,672	59,703
31	12,196	14,458	15,655	17,539	19,281	21,434	41,422	44,985	48,232	52,191	55,003	61,098
32	12,811	15,134	16,362	18,291	20,072	22,271	42,585	46,194	49,480	53,486	56,328	62,487
33	13,431	15,815	17,074	19,047	20,867	23,110	43,745	47,400	50,725	54,776	57,648	63,870
34	14,057	16,501	17,789	19,806	21,664	23,952	44,903	48,602	51,966	56,061	58,964	65,247
35	14,688	17,192	18,509	20,569	22,465	24,797	46,059	49,802	53,203	57,342	60,275	66,619
36	15,324	17,887	19,233	21,336	23,269	25,643	47,212	50,998	54,437	58,619	61,581	67,985
37	15,965	18,586	19,960	22,106	24,075	26,492	48,363	52,192	55,668	59,893	62,883	69,346
38	16,611	19,289	20,691	22,878	24,884	27,343	49,513	53,384	56,896	61,162	64,181	70,703
39	17,262	19,996	21,426	23,654	25,695	28,196	50,660	54,572	58,120	62,428	65,476	72,055
40	17,916	20,707	22,164	24,433	26,509	29,051	51,805	55,758	59,342	63,691	66,766	73,402
41	18,575	21,421	22,906	25,215	27,326	29,907	52,949	56,942	60,561	64,950	68,053	74,745
42	19,239	22,138	23,650	25,999	28,144	30,765	54,090	58,124	61,777	66,206	69,336	76,084
43	19,906	22,859	24,398	26,785	28,965	31,625	55,230	59,303	62,990	67,459	70,616	77,419
44	20,576	23,584	25,148	27,575	29,787	32,487	56,369	60,481	64,201	68,710	71,893	78,750
45	21,251	24,311	25,901	28,366	30,612	33,350	57,505	61,656	65,410	69,957	73,166	80,077
46	21,929	25,041	26,657	29,160	31,439	34,215	58,641	62,830	66,617	71,201	74,437	81,400
47	22,610	25,775	27,416	29,956	32,268	35,081	59,774	64,001	67,821	72,443	75,704	82,720
48	23,295	26,511	28,177	30,755	33,098	35,949	60,907	65,171	69,023	73,683	76,969	84,037
49	23,983	27,249	28,941	31,555	33,930	36,818	62,038	66,339	70,222	74,919	78,231	85,351
50	24,674	27,991	29,707	32,357	34,764	37,689	63,167	67,505	71,420	76,154	79,490	86,661

ν	$1-\alpha$											
	0,001	0,005	0,010	0,025	0,050	0,100	0,900	0,950	0,975	0,990	0,995	0,999
51	25,368	28,735	30,475	33,162	35,600	38,560	64,295	68,669	72,616	77,386	80,747	87,968
52	26,065	29,481	31,246	33,968	36,437	39,433	65,422	69,832	73,810	78,616	82,001	89,272
53	26,765	30,230	32,018	34,776	37,276	40,308	66,548	70,993	75,002	79,843	83,253	90,573
54	27,468	30,981	32,793	35,586	38,116	41,183	67,673	72,153	76,192	81,069	84,502	91,872
55	28,173	31,735	33,570	36,398	38,958	42,060	68,796	73,311	77,380	82,292	85,749	93,168
56	28,881	32,490	34,350	37,212	39,801	42,937	69,919	74,468	78,567	83,513	86,994	94,461
57	29,592	33,248	35,131	38,027	40,646	43,816	71,040	75,624	79,752	84,733	88,236	95,751
58	30,305	34,008	35,913	38,844	41,492	44,696	72,160	76,778	80,936	85,950	89,477	97,039
59	31,020	34,770	36,698	39,662	42,339	45,577	73,279	77,931	82,117	87,166	90,715	98,324
60	31,738	35,534	37,485	40,482	43,188	46,459	74,397	79,082	83,298	88,379	91,952	99,607
61	32,459	36,300	38,273	41,303	44,038	47,342	75,514	80,232	84,476	89,591	93,186	100,888
62	33,181	37,068	39,063	42,126	44,889	48,226	76,630	81,381	85,654	90,802	94,419	102,166
63	33,906	37,838	39,855	42,950	45,741	49,111	77,745	82,529	86,830	92,010	95,649	103,442
64	34,633	38,610	40,649	43,776	46,595	49,996	78,860	83,675	88,004	93,217	96,878	104,716
65	35,362	39,383	41,444	44,603	47,450	50,883	79,973	84,821	89,177	94,422	98,105	105,988
66	36,093	40,158	42,240	45,431	48,305	51,770	81,085	85,965	90,349	95,626	99,330	107,258
67	36,826	40,935	43,038	46,261	49,162	52,659	82,197	87,108	91,519	96,828	100,554	108,526
68	37,561	41,713	43,838	47,092	50,020	53,548	83,308	88,250	92,689	98,028	101,776	109,791
69	38,298	42,494	44,639	47,924	50,879	54,438	84,418	89,391	93,856	99,228	102,996	111,055
70	39,036	43,275	45,442	48,758	51,739	55,329	85,527	90,531	95,023	100,425	104,215	112,317
75	42,757	47,206	49,475	52,942	56,054	59,795	91,061	96,217	100,839	106,393	110,286	118,599
80	46,520	51,172	53,540	57,153	60,391	64,278	96,578	101,879	106,629	112,329	116,321	124,839
85	50,320	55,170	57,634	61,389	64,749	68,777	102,079	107,522	112,393	118,236	122,325	131,041
90	54,155	59,196	61,754	65,647	69,126	73,291	107,565	113,145	118,136	124,116	128,299	137,208
95	58,022	63,250	65,898	69,925	73,520	77,818	113,038	118,752	123,858	129,973	134,247	143,344
100	61,918	67,328	70,065	74,222	77,929	82,358	118,498	124,342	129,561	135,807	140,169	149,449

F-Verteilung:



In Abhängigkeit von ν_1 und ν_2 sind die Werte x der F-verteilten Zufallsvariablen tabelliert, bei denen die Verteilungsfunktion $(1-\alpha)$ die Werte 0,90; 0,95; 0,975; 0,99 und 0,995 annimmt (Quelle: J. Schwarze, Grundlagen der Statistik II, 1993, S. 301 - 305)

$1-\alpha = 0,900$

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	14	16	18	20	30	40	50	100	∞
1	39,86	49,50	53,59	55,83	57,24	58,20	59,14	59,69	60,12	60,46	60,98	61,36	61,64	61,86	62,04	62,58	62,85	63,01	63,33	63,33
2	8,526	9,000	9,162	9,244	9,293	9,326	9,350	9,367	9,381	9,392	9,409	9,421	9,430	9,437	9,442	9,459	9,467	9,472	9,482	9,491
3	5,538	5,462	5,390	5,343	5,309	5,285	5,266	5,251	5,240	5,230	5,215	5,204	5,196	5,189	5,184	5,167	5,159	5,154	5,143	5,134
4	4,545	4,325	4,191	4,107	4,051	4,010	3,979	3,955	3,936	3,920	3,895	3,877	3,864	3,853	3,844	3,817	3,803	3,795	3,777	3,761
5	4,060	3,780	3,619	3,520	3,453	3,404	3,368	3,339	3,316	3,297	3,268	3,247	3,230	3,217	3,206	3,174	3,157	3,147	3,126	3,105
6	3,776	3,463	3,289	3,181	3,107	3,055	3,014	2,983	2,958	2,937	2,905	2,881	2,863	2,848	2,836	2,800	2,781	2,770	2,746	2,722
7	3,589	3,257	3,074	2,961	2,883	2,827	2,785	2,752	2,725	2,703	2,668	2,643	2,623	2,607	2,595	2,555	2,535	2,523	2,497	2,471
8	3,458	3,113	2,924	2,806	2,726	2,668	2,624	2,589	2,561	2,538	2,502	2,475	2,454	2,438	2,425	2,383	2,361	2,348	2,321	2,293
9	3,360	3,006	2,813	2,693	2,611	2,551	2,505	2,469	2,440	2,416	2,379	2,351	2,329	2,312	2,298	2,255	2,232	2,218	2,189	2,159
10	3,285	2,924	2,728	2,605	2,522	2,461	2,414	2,377	2,347	2,323	2,284	2,255	2,233	2,215	2,201	2,155	2,132	2,117	2,087	2,055
11	3,225	2,859	2,660	2,536	2,451	2,389	2,342	2,304	2,274	2,248	2,209	2,179	2,156	2,138	2,123	2,076	2,052	2,036	2,005	1,972
12	3,177	2,807	2,606	2,480	2,394	2,331	2,283	2,245	2,214	2,188	2,147	2,117	2,094	2,075	2,060	2,011	1,986	1,970	1,938	1,904
13	3,136	2,763	2,560	2,434	2,347	2,283	2,234	2,195	2,164	2,138	2,097	2,066	2,042	2,023	2,007	1,958	1,931	1,915	1,882	1,846
14	3,102	2,726	2,522	2,395	2,307	2,243	2,193	2,154	2,122	2,095	2,054	2,022	1,998	1,978	1,962	1,912	1,885	1,869	1,834	1,797
15	3,073	2,695	2,490	2,361	2,273	2,208	2,158	2,119	2,086	2,059	2,017	1,985	1,961	1,941	1,924	1,873	1,845	1,828	1,793	1,755
16	3,048	2,668	2,462	2,333	2,244	2,178	2,128	2,088	2,055	2,028	1,985	1,953	1,928	1,908	1,891	1,839	1,811	1,793	1,757	1,718
17	3,026	2,645	2,437	2,308	2,218	2,152	2,102	2,061	2,028	2,001	1,958	1,925	1,900	1,879	1,862	1,809	1,781	1,763	1,726	1,686
18	3,007	2,624	2,416	2,286	2,196	2,130	2,079	2,038	2,005	1,977	1,933	1,900	1,875	1,854	1,837	1,783	1,754	1,736	1,698	1,657
19	2,990	2,605	2,397	2,266	2,176	2,109	2,058	2,017	1,984	1,956	1,912	1,878	1,852	1,831	1,814	1,759	1,730	1,711	1,673	1,631
20	2,975	2,589	2,380	2,249	2,158	2,091	2,040	1,999	1,965	1,937	1,892	1,859	1,833	1,811	1,794	1,738	1,708	1,690	1,650	1,607
21	2,961	2,574	2,365	2,233	2,142	2,075	2,023	1,982	1,948	1,920	1,875	1,841	1,815	1,793	1,776	1,719	1,689	1,670	1,630	1,586
22	2,949	2,561	2,351	2,219	2,128	2,060	2,008	1,967	1,933	1,904	1,859	1,825	1,798	1,777	1,759	1,702	1,671	1,652	1,611	1,567
23	2,937	2,549	2,339	2,207	2,115	2,047	1,995	1,953	1,919	1,890	1,845	1,811	1,784	1,762	1,744	1,686	1,655	1,636	1,594	1,549
24	2,927	2,538	2,327	2,195	2,103	2,035	1,983	1,941	1,906	1,877	1,832	1,797	1,770	1,748	1,730	1,672	1,641	1,621	1,579	1,533
25	2,918	2,528	2,317	2,184	2,092	2,024	1,971	1,929	1,895	1,866	1,820	1,785	1,758	1,736	1,718	1,659	1,627	1,607	1,565	1,518
26	2,909	2,519	2,307	2,174	2,082	2,014	1,961	1,919	1,884	1,855	1,809	1,774	1,747	1,724	1,706	1,647	1,615	1,594	1,551	1,504
27	2,901	2,511	2,299	2,165	2,073	2,005	1,951	1,909	1,874	1,845	1,799	1,764	1,736	1,714	1,695	1,636	1,603	1,583	1,539	1,491
28	2,894	2,503	2,291	2,157	2,064	1,996	1,943	1,900	1,865	1,836	1,790	1,754	1,726	1,704	1,685	1,625	1,592	1,572	1,528	1,478
29	2,887	2,495	2,283	2,149	2,057	1,988	1,934	1,892	1,857	1,827	1,781	1,745	1,717	1,695	1,676	1,616	1,583	1,562	1,517	1,467
30	2,881	2,489	2,276	2,142	2,049	1,980	1,927	1,884	1,849	1,819	1,773	1,737	1,709	1,686	1,667	1,606	1,573	1,552	1,507	1,456
40	2,835	2,440	2,226	2,091	1,997	1,927	1,872	1,829	1,793	1,763	1,715	1,678	1,649	1,625	1,605	1,541	1,506	1,483	1,434	1,377
50	2,809	2,412	2,197	2,061	1,966	1,895	1,840	1,796	1,760	1,729	1,680	1,643	1,613	1,588	1,568	1,502	1,465	1,441	1,388	1,327
60	2,791	2,393	2,177	2,041	1,946	1,875	1,819	1,775	1,738	1,707	1,657	1,619	1,589	1,564	1,543	1,476	1,437	1,413	1,358	1,291
70	2,779	2,380	2,164	2,027	1,931	1,860	1,804	1,760	1,723	1,691	1,641	1,603	1,572	1,547	1,526	1,457	1,418	1,392	1,335	1,265
80	2,769	2,370	2,153	2,016	1,921	1,849	1,793	1,748	1,711	1,680	1,629	1,590	1,559	1,534	1,513	1,443	1,403	1,377	1,318	1,245
90	2,762	2,362	2,146	2,008	1,912	1,841	1,785	1,739	1,702	1,670	1,620	1,581	1,550	1,524	1,503	1,432	1,391	1,365	1,304	1,228
100	2,756	2,356	2,139	2,002	1,906	1,834	1,778	1,732	1,695	1,663	1,612	1,573	1,542	1,516	1,494	1,423	1,382	1,355	1,293	1,214
∞	2,706	2,303	2,084	1,945	1,847	1,774	1,717	1,670	1,632	1,599	1,546	1,505	1,471	1,444	1,421	1,342	1,295	1,263	1,185	1,000

	$1-\alpha = 0{,}950$																			
V_2	V_1	1	2	3	4	5	6	7	8	9	10	12	14	16	18	20	30	40	50	100	∞
1		161,4	199,5	215,8	224,7	230,4	234,2	237,0	239,1	240,8	242,1	244,2	245,6	246,8	247,6	248,3	250,4	251,5	252,1	253,4	254,3
2		18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,41	19,42	19,43	19,44	19,45	19,46	19,47	19,48	19,49	19,50
3		10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,71	8,69	8,67	8,66	8,62	8,59	8,58	8,54	8,53
4		7,709	6,944	6,591	6,388	6,256	6,163	6,094	6,041	5,999	5,964	5,912	5,873	5,844	5,821	5,803	5,746	5,717	5,699	5,658	5,628
5		6,608	5,786	5,409	5,192	5,050	4,950	4,876	4,818	4,772	4,735	4,678	4,636	4,604	4,579	4,558	4,496	4,464	4,444	4,403	4,365
6		5,987	5,143	4,757	4,534	4,387	4,284	4,206	4,147	4,099	4,060	4,000	3,956	3,922	3,896	3,874	3,808	3,773	3,753	3,711	3,669
7		5,591	4,737	4,347	4,120	3,971	3,866	3,787	3,726	3,677	3,636	3,574	3,529	3,494	3,467	3,444	3,375	3,340	3,318	3,274	3,230
8		5,318	4,459	4,066	3,838	3,687	3,581	3,500	3,438	3,388	3,347	3,284	3,237	3,202	3,173	3,150	3,079	3,042	3,020	2,974	2,928
9		5,117	4,256	3,863	3,633	3,482	3,374	3,293	3,230	3,179	3,137	3,073	3,025	2,989	2,960	2,936	2,864	2,826	2,803	2,755	2,707
10		4,965	4,103	3,708	3,478	3,326	3,217	3,135	3,072	3,020	2,978	2,913	2,865	2,828	2,798	2,774	2,699	2,661	2,637	2,588	2,538
11		4,844	3,982	3,587	3,357	3,204	3,095	3,012	2,948	2,896	2,854	2,788	2,739	2,701	2,671	2,646	2,570	2,531	2,507	2,456	2,404
12		4,747	3,885	3,490	3,259	3,106	2,996	2,913	2,849	2,796	2,753	2,687	2,637	2,599	2,568	2,544	2,466	2,426	2,401	2,350	2,296
13		4,667	3,806	3,411	3,179	3,025	2,915	2,832	2,767	2,714	2,671	2,604	2,554	2,515	2,484	2,459	2,380	2,339	2,314	2,261	2,206
14		4,600	3,739	3,344	3,112	2,958	2,848	2,764	2,699	2,646	2,602	2,534	2,484	2,445	2,413	2,388	2,308	2,266	2,240	2,187	2,131
15		4,543	3,682	3,287	3,056	2,901	2,790	2,707	2,641	2,588	2,544	2,475	2,424	2,385	2,353	2,328	2,247	2,204	2,178	2,123	2,066
16		4,494	3,634	3,239	3,007	2,852	2,741	2,657	2,591	2,538	2,494	2,425	2,373	2,333	2,302	2,276	2,194	2,151	2,124	2,068	2,010
17		4,451	3,592	3,197	2,965	2,810	2,699	2,614	2,548	2,494	2,450	2,381	2,329	2,289	2,257	2,230	2,148	2,104	2,077	2,020	1,960
18		4,414	3,555	3,160	2,928	2,773	2,661	2,577	2,510	2,456	2,412	2,342	2,290	2,250	2,217	2,191	2,107	2,063	2,035	1,978	1,917
19		4,381	3,522	3,127	2,895	2,740	2,628	2,544	2,477	2,423	2,378	2,308	2,256	2,215	2,182	2,155	2,071	2,026	1,999	1,940	1,878
20		4,351	3,493	3,098	2,866	2,711	2,599	2,514	2,447	2,393	2,348	2,278	2,225	2,184	2,151	2,124	2,039	1,994	1,966	1,907	1,843
21		4,325	3,467	3,072	2,840	2,685	2,573	2,488	2,420	2,366	2,321	2,250	2,197	2,156	2,123	2,096	2,010	1,965	1,936	1,876	1,812
22		4,301	3,443	3,049	2,817	2,661	2,549	2,464	2,396	2,342	2,297	2,226	2,173	2,131	2,098	2,071	1,984	1,938	1,909	1,849	1,783
23		4,279	3,422	3,028	2,796	2,640	2,528	2,442	2,375	2,320	2,275	2,204	2,150	2,109	2,075	2,048	1,961	1,914	1,885	1,823	1,757
24		4,260	3,403	3,009	2,776	2,621	2,508	2,423	2,355	2,300	2,255	2,183	2,130	2,088	2,054	2,027	1,939	1,892	1,863	1,800	1,733
25		4,242	3,385	2,991	2,759	2,603	2,490	2,405	2,337	2,282	2,236	2,165	2,111	2,069	2,035	2,007	1,919	1,872	1,842	1,779	1,711
26		4,225	3,369	2,975	2,743	2,587	2,474	2,388	2,321	2,265	2,220	2,148	2,094	2,052	2,018	1,990	1,901	1,853	1,823	1,760	1,691
27		4,210	3,354	2,960	2,728	2,572	2,459	2,373	2,305	2,250	2,204	2,132	2,078	2,036	2,002	1,974	1,884	1,836	1,806	1,742	1,672
28		4,196	3,340	2,947	2,714	2,558	2,445	2,359	2,291	2,236	2,190	2,118	2,064	2,021	1,987	1,959	1,869	1,820	1,790	1,725	1,654
29		4,183	3,327	2,934	2,701	2,545	2,432	2,346	2,278	2,223	2,177	2,104	2,050	2,007	1,973	1,945	1,854	1,806	1,775	1,710	1,638
30		4,171	3,316	2,922	2,690	2,534	2,420	2,334	2,266	2,211	2,165	2,092	2,037	1,995	1,960	1,932	1,841	1,792	1,761	1,695	1,622
40		4,085	3,232	2,839	2,606	2,449	2,336	2,249	2,180	2,124	2,077	2,003	1,948	1,904	1,868	1,839	1,744	1,693	1,660	1,589	1,509
50		4,034	3,183	2,790	2,557	2,400	2,286	2,199	2,130	2,073	2,026	1,952	1,895	1,850	1,814	1,784	1,687	1,634	1,599	1,525	1,438
60		4,001	3,150	2,758	2,525	2,368	2,254	2,166	2,097	2,040	1,993	1,917	1,860	1,815	1,778	1,748	1,649	1,594	1,559	1,481	1,389
70		3,978	3,127	2,735	2,503	2,346	2,231	2,143	2,074	2,017	1,969	1,893	1,836	1,790	1,753	1,722	1,622	1,566	1,530	1,450	1,353
80		3,960	3,110	2,719	2,486	2,329	2,214	2,126	2,056	1,999	1,951	1,875	1,817	1,772	1,734	1,703	1,602	1,545	1,508	1,426	1,325
90		3,947	3,097	2,706	2,473	2,316	2,201	2,113	2,043	1,986	1,938	1,861	1,803	1,757	1,720	1,688	1,586	1,528	1,491	1,407	1,302
100		3,936	3,087	2,695	2,463	2,305	2,191	2,102	2,032	1,975	1,927	1,850	1,792	1,746	1,708	1,676	1,573	1,515	1,477	1,392	1,283
∞		3,841	2,996	2,605	2,372	2,214	2,099	2,010	1,938	1,880	1,831	1,752	1,692	1,644	1,604	1,571	1,459	1,394	1,350	1,243	1,000

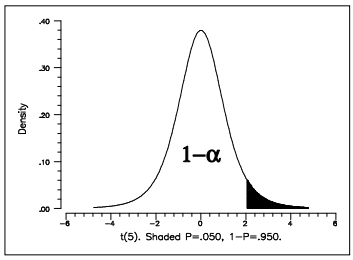
$1-\alpha = 0{,}975$

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	14	16	18	20	30	40	50	100	∞
1	647,5	799,5	864,3	899,8	922,0	937,3	948,5	956,9	963,5	968,9	977,0	982,8	987,2	990,6	993,4	1002	1006	1008	1014	1018
2	38,51	39,00	39,17	39,25	39,30	39,33	39,36	39,37	39,39	39,40	39,41	39,43	39,44	39,44	39,45	39,46	39,47	39,48	39,98	39,50
3	17,44	16,04	15,44	15,10	14,88	14,73	14,62	14,54	14,47	14,42	14,34	14,28	14,23	14,20	14,17	14,08	14,04	14,01	13,96	13,90
4	12,22	10,65	9,979	9,605	9,364	9,197	9,074	8,980	8,905	8,844	8,751	8,684	8,633	8,592	8,560	8,461	8,411	8,381	8,319	8,257
5	10,01	8,43	7,764	7,388	7,146	6,978	6,853	6,757	6,681	6,619	6,525	6,456	6,403	6,362	6,329	6,227	6,175	6,144	6,080	6,015
6	8,813	7,260	6,599	6,227	5,988	5,820	5,695	5,600	5,523	5,461	5,366	5,297	5,244	5,202	5,168	5,065	5,012	4,980	4,912	4,849
7	8,073	6,542	5,890	5,523	5,285	5,119	4,995	4,899	4,823	4,761	4,666	4,596	4,543	4,501	4,467	4,362	4,309	4,276	4,208	4,142
8	7,571	6,059	5,416	5,053	4,817	4,651	4,528	4,433	4,357	4,295	4,200	4,130	4,076	4,034	3,999	3,894	3,839	3,807	3,738	3,670
9	7,209	5,715	5,078	4,718	4,484	4,320	4,197	4,102	4,026	3,964	3,868	3,798	3,744	3,701	3,667	3,560	3,505	3,472	3,403	3,333
10	6,937	5,456	4,826	4,468	4,236	4,072	3,950	3,855	3,779	3,717	3,621	3,550	3,496	3,453	3,418	3,311	3,255	3,221	3,151	3,080
11	6,724	5,256	4,630	4,275	4,044	3,881	3,759	3,664	3,588	3,526	3,430	3,359	3,304	3,261	3,226	3,117	3,061	3,027	2,956	2,883
12	6,554	5,096	4,474	4,121	3,891	3,728	3,606	3,512	3,436	3,374	3,277	3,206	3,152	3,108	3,073	2,963	2,906	2,871	2,799	2,725
13	6,414	4,965	4,347	3,996	3,767	3,604	3,483	3,388	3,312	3,250	3,153	3,082	3,027	2,983	2,948	2,837	2,780	2,744	2,671	2,595
14	6,298	4,857	4,242	3,892	3,663	3,501	3,380	3,285	3,209	3,147	3,050	2,979	2,923	2,879	2,844	2,732	2,674	2,638	2,564	2,487
15	6,200	4,765	4,153	3,804	3,576	3,415	3,293	3,199	3,123	3,060	2,963	2,891	2,836	2,792	2,756	2,644	2,585	2,549	2,474	2,395
16	6,115	4,687	4,077	3,729	3,502	3,341	3,219	3,125	3,049	2,986	2,889	2,817	2,761	2,717	2,681	2,568	2,508	2,472	2,396	2,316
17	6,042	4,619	4,011	3,665	3,438	3,277	3,156	3,061	2,985	2,922	2,825	2,753	2,697	2,652	2,616	2,502	2,442	2,405	2,328	2,247
18	5,978	4,560	3,954	3,608	3,382	3,221	3,100	3,005	2,929	2,866	2,769	2,696	2,640	2,596	2,559	2,444	2,384	2,347	2,269	2,187
19	5,922	4,508	3,903	3,559	3,333	3,172	3,051	2,956	2,880	2,817	2,720	2,647	2,591	2,546	2,509	2,394	2,333	2,295	2,217	2,133
20	5,871	4,461	3,859	3,515	3,289	3,128	3,007	2,913	2,837	2,774	2,676	2,603	2,547	2,501	2,464	2,349	2,287	2,249	2,170	2,085
21	5,827	4,420	3,819	3,475	3,250	3,089	2,969	2,874	2,798	2,735	2,637	2,564	2,507	2,462	2,425	2,308	2,246	2,208	2,128	2,042
22	5,786	4,383	3,783	3,440	3,215	3,055	2,934	2,839	2,763	2,700	2,602	2,528	2,472	2,426	2,389	2,272	2,210	2,171	2,090	2,003
23	5,750	4,349	3,750	3,408	3,183	3,023	2,902	2,808	2,731	2,668	2,570	2,497	2,440	2,394	2,357	2,239	2,176	2,137	2,056	1,968
24	5,717	4,319	3,721	3,379	3,155	2,995	2,874	2,779	2,703	2,640	2,541	2,468	2,411	2,365	2,327	2,209	2,146	2,107	2,024	1,935
25	5,686	4,291	3,694	3,353	3,129	2,968	2,848	2,753	2,677	2,613	2,515	2,441	2,384	2,338	2,300	2,182	2,118	2,079	1,996	1,906
26	5,659	4,265	3,670	3,329	3,105	2,945	2,824	2,729	2,653	2,590	2,491	2,417	2,360	2,314	2,276	2,157	2,093	2,053	1,969	1,878
27	5,633	4,242	3,647	3,307	3,083	2,923	2,802	2,707	2,631	2,568	2,469	2,395	2,337	2,291	2,253	2,133	2,069	2,029	1,945	1,853
28	5,610	4,221	3,626	3,286	3,063	2,903	2,782	2,687	2,611	2,547	2,448	2,374	2,317	2,270	2,232	2,112	2,048	2,007	1,922	1,829
29	5,588	4,201	3,607	3,267	3,044	2,884	2,763	2,669	2,592	2,529	2,430	2,355	2,298	2,251	2,213	2,092	2,028	1,987	1,901	1,807
30	5,568	4,182	3,589	3,250	3,026	2,867	2,746	2,651	2,575	2,511	2,412	2,338	2,280	2,233	2,195	2,074	2,009	1,968	1,882	1,787
40	5,424	4,051	3,463	3,126	2,904	2,744	2,624	2,529	2,452	2,388	2,288	2,213	2,154	2,107	2,068	1,943	1,875	1,832	1,741	1,637
50	5,340	3,975	3,390	3,054	2,833	2,673	2,553	2,458	2,381	2,317	2,216	2,140	2,081	2,033	1,993	1,866	1,796	1,752	1,656	1,545
60	5,286	3,925	3,342	3,008	2,786	2,627	2,507	2,412	2,334	2,270	2,169	2,093	2,033	1,985	1,944	1,815	1,744	1,699	1,599	1,482
70	5,247	3,890	3,309	2,975	2,754	2,595	2,474	2,379	2,302	2,237	2,136	2,059	1,999	1,950	1,910	1,779	1,707	1,660	1,558	1,436
80	5,218	3,864	3,284	2,950	2,730	2,571	2,450	2,355	2,277	2,213	2,111	2,035	1,974	1,925	1,884	1,752	1,679	1,632	1,527	1,400
90	5,196	3,844	3,265	2,932	2,711	2,552	2,431	2,336	2,259	2,194	2,092	2,015	1,955	1,905	1,864	1,731	1,657	1,610	1,503	1,371
100	5,179	3,828	3,249	2,917	2,696	2,537	2,417	2,321	2,244	2,179	2,077	2,000	1,939	1,890	1,849	1,715	1,640	1,592	1,483	1,347
∞	5,024	3,689	3,116	2,786	2,567	2,408	2,288	2,192	2,114	2,048	1,945	1,866	1,803	1,751	1,708	1,566	1,484	1,428	1,296	1,000

	$1-\alpha = 0{,}990$																			
v_2 \ v_1	1	2	3	4	5	6	7	8	9	10	12	14	16	18	20	30	40	50	100	∞
1	4052	5000	5403	5625	5764	5859	5929	5981	6023	6056	6107	6143	6170	6192	6209	6261	6287	6303	6334	6366
2	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,75	99,78	99,80	99,83	99,86	99,88	99,89	99,90	99,93	99,95	99,96	99,98	99,50
3	34,11	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,34	27,23	27,05	26,92	26,83	26,75	26,69	26,50	26,41	26,35	26,24	26,13
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55	14,37	14,25	14,15	14,08	14,02	13,84	13,75	13,69	13,58	13,46
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,888	9,770	9,680	9,610	9,553	9,379	9,291	9,238	9,130	9,020
6	13,74	10,92	9,779	9,148	8,746	8,466	8,260	8,102	7,976	7,874	7,718	7,605	7,519	7,451	7,396	7,229	7,143	7,091	6,987	6,880
7	12,25	9,55	8,451	7,847	7,460	7,191	6,993	6,840	6,719	6,620	6,469	6,359	6,275	6,209	6,155	5,992	5,908	5,858	5,751	5,650
8	11,26	8,65	7,591	7,006	6,632	6,371	6,178	6,029	5,911	5,814	5,667	5,559	5,477	5,412	5,359	5,198	5,116	5,065	4,961	4,859
9	10,56	8,02	6,992	6,422	6,057	5,802	5,613	5,467	5,351	5,257	5,111	5,005	4,924	4,860	4,808	4,649	4,567	4,516	4,414	4,311
10	10,04	7,56	6,552	5,994	5,636	5,386	5,200	5,057	4,942	4,849	4,706	4,601	4,520	4,457	4,405	4,246	4,165	4,115	4,013	3,909
11	9,646	7,206	6,217	5,668	5,316	5,069	4,886	4,744	4,631	4,539	4,397	4,293	4,213	4,150	4,099	3,941	3,859	3,809	3,707	3,602
12	9,330	6,927	5,952	5,412	5,064	4,820	4,639	4,499	4,387	4,296	4,155	4,052	3,972	3,909	3,858	3,701	3,619	3,569	3,466	3,361
13	9,074	6,701	5,739	5,205	4,862	4,620	4,441	4,302	4,191	4,100	3,960	3,857	3,778	3,715	3,664	3,507	3,425	3,375	3,272	3,165
14	8,862	6,515	5,564	5,035	4,695	4,456	4,278	4,140	4,030	3,939	3,800	3,697	3,619	3,556	3,505	3,347	3,266	3,215	3,112	3,004
15	8,683	6,359	5,417	4,893	4,556	4,318	4,141	4,004	3,895	3,805	3,666	3,564	3,485	3,423	3,372	3,214	3,132	3,081	2,977	2,868
16	8,531	6,226	5,292	4,773	4,437	4,202	4,026	3,890	3,780	3,691	3,553	3,451	3,372	3,310	3,259	3,101	3,018	2,967	2,863	2,753
17	8,400	6,112	5,185	4,669	4,336	4,101	3,927	3,791	3,682	3,593	3,455	3,353	3,275	3,212	3,161	3,003	2,920	2,869	2,764	2,653
18	8,285	6,013	5,092	4,579	4,248	4,015	3,841	3,705	3,597	3,508	3,371	3,269	3,190	3,128	3,077	2,918	2,835	2,784	2,678	2,566
19	8,185	5,926	5,010	4,500	4,171	3,939	3,765	3,631	3,522	3,434	3,297	3,195	3,116	3,054	3,003	2,844	2,761	2,709	2,602	2,489
20	8,096	5,849	4,938	4,431	4,103	3,871	3,699	3,564	3,457	3,368	3,231	3,130	3,051	2,989	2,938	2,778	2,695	2,643	2,535	2,421
21	8,016	5,780	4,874	4,369	4,042	3,812	3,640	3,506	3,398	3,310	3,173	3,071	2,993	2,931	2,880	2,720	2,636	2,584	2,475	2,360
22	7,945	5,719	4,817	4,313	3,988	3,758	3,587	3,453	3,346	3,258	3,121	3,019	2,941	2,879	2,827	2,667	2,583	2,531	2,422	2,305
23	7,881	5,664	4,765	4,264	3,939	3,710	3,539	3,406	3,299	3,211	3,074	2,973	2,894	2,832	2,780	2,620	2,535	2,483	2,373	2,256
24	7,823	5,614	4,718	4,218	3,895	3,667	3,496	3,363	3,256	3,168	3,032	2,930	2,852	2,789	2,738	2,577	2,492	2,439	2,329	2,211
25	7,770	5,568	4,675	4,177	3,855	3,627	3,457	3,324	3,217	3,129	2,993	2,892	2,813	2,751	2,699	2,538	2,453	2,400	2,289	2,169
26	7,721	5,526	4,636	4,140	3,818	3,591	3,421	3,288	3,182	3,094	2,958	2,857	2,778	2,715	2,664	2,503	2,417	2,364	2,252	2,131
27	7,677	5,488	4,601	4,106	3,785	3,558	3,388	3,256	3,149	3,062	2,926	2,824	2,746	2,683	2,632	2,470	2,384	2,330	2,218	2,097
28	7,636	5,453	4,568	4,074	3,754	3,528	3,358	3,226	3,120	3,032	2,896	2,795	2,716	2,653	2,602	2,440	2,354	2,300	2,187	2,064
29	7,598	5,420	4,538	4,045	3,725	3,499	3,330	3,198	3,092	3,005	2,868	2,767	2,689	2,626	2,574	2,412	2,325	2,271	2,158	2,034
30	7,562	5,390	4,510	4,018	3,699	3,473	3,304	3,173	3,067	2,979	2,843	2,742	2,663	2,600	2,549	2,386	2,299	2,245	2,131	2,006
40	7,314	5,179	4,312	3,828	3,514	3,291	3,124	2,993	2,888	2,801	2,665	2,563	2,484	2,421	2,369	2,203	2,114	2,058	1,938	1,805
50	7,170	5,056	4,199	3,720	3,408	3,186	3,020	2,890	2,785	2,698	2,563	2,461	2,382	2,318	2,265	2,098	2,007	1,949	1,825	1,683
60	7,077	4,977	4,126	3,649	3,339	3,119	2,953	2,823	2,718	2,632	2,496	2,394	2,315	2,251	2,198	2,028	1,936	1,877	1,749	1,601
70	7,011	4,922	4,074	3,600	3,291	3,071	2,906	2,777	2,672	2,585	2,450	2,348	2,268	2,204	2,150	1,980	1,886	1,826	1,695	1,540
80	6,962	4,881	4,036	3,563	3,255	3,036	2,871	2,742	2,637	2,551	2,415	2,313	2,233	2,169	2,115	1,944	1,849	1,788	1,655	1,494
90	6,925	4,849	4,007	3,535	3,228	3,009	2,844	2,715	2,611	2,524	2,389	2,286	2,206	2,142	2,088	1,916	1,820	1,759	1,623	1,457
100	6,895	4,824	3,983	3,513	3,206	2,988	2,823	2,694	2,590	2,503	2,368	2,265	2,185	2,120	2,067	1,893	1,797	1,735	1,598	1,427
∞	6,635	4,605	3,782	3,319	3,017	2,802	2,639	2,511	2,407	2,321	2,185	2,082	2,000	1,934	1,878	1,696	1,592	1,523	1,358	1,000

	$1-\alpha = 0{,}995$																			
ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	14	16	18	20	30	40	50	100	∞
1	16210	20000	21615	22500	23056	23437	23715	23926	24091	24225	24427	24572	24682	24768	24836	25044	25149	25211	25338	25465
2	198,5	199,0	199,3	199,5	199,6	199,7	199,7	199,8	199,8	199,8	199,8	199,9	199,9	199,9	199,9	199,9	200,0	200,0	200,0	199,5
3	55,55	49,80	47,46	46,19	45,39	44,84	44,43	44,12	43,88	43,68	43,38	43,17	43,00	42,88	42,77	42,46	42,30	42,21	42,02	41,83
4	31,33	26,28	24,26	23,15	22,46	21,97	21,62	21,35	21,14	20,97	20,70	20,51	20,37	20,26	20,17	19,89	19,75	19,67	19,50	19,32
5	22,78	18,31	16,53	15,56	14,94	14,51	14,20	13,96	13,77	13,62	13,38	13,21	13,09	12,98	12,90	12,66	12,53	12,45	12,30	12,14
6	18,635	14,544	12,916	12,028	11,464	11,073	10,786	10,566	10,391	10,250	10,034	9,877	9,758	9,664	9,589	9,358	9,241	9,170	9,026	8,879
7	16,236	12,404	10,882	10,050	9,522	9,155	8,885	8,678	8,514	8,380	8,176	8,028	7,915	7,826	7,754	7,534	7,422	7,354	7,210	7,076
8	14,688	11,042	9,596	8,805	8,302	7,952	7,694	7,496	7,339	7,211	7,015	6,872	6,763	6,678	6,608	6,396	6,288	6,222	6,084	5,951
9	13,614	10,107	8,717	7,956	7,471	7,134	6,885	6,693	6,541	6,417	6,227	6,089	5,983	5,899	5,832	5,625	5,519	5,452	5,321	5,188
10	12,826	9,427	8,081	7,343	6,872	6,545	6,302	6,116	5,968	5,847	5,661	5,526	5,422	5,340	5,274	5,071	4,965	4,901	4,771	4,639
11	12,226	8,912	7,600	6,881	6,422	6,102	5,865	5,682	5,537	5,418	5,236	5,103	5,001	4,921	4,855	4,654	4,550	4,487	4,358	4,226
12	11,754	8,510	7,226	6,521	6,071	5,757	5,525	5,345	5,202	5,085	4,906	4,774	4,674	4,594	4,530	4,331	4,228	4,165	4,036	3,904
13	11,373	8,186	6,926	6,233	5,791	5,482	5,253	5,076	4,935	4,820	4,643	4,513	4,413	4,334	4,270	4,073	3,970	3,908	3,779	3,647
14	11,060	7,922	6,680	5,998	5,562	5,257	5,031	4,856	4,717	4,603	4,428	4,299	4,200	4,122	4,058	3,862	3,760	3,697	3,569	3,436
15	10,798	7,701	6,476	5,803	5,372	5,071	4,847	4,674	4,536	4,423	4,250	4,122	4,024	3,946	3,882	3,687	3,585	3,522	3,394	3,260
16	10,575	7,514	6,303	5,638	5,212	4,913	4,692	4,521	4,384	4,272	4,099	3,972	3,875	3,797	3,734	3,539	3,437	3,375	3,246	3,112
17	10,384	7,354	6,155	5,497	5,075	4,779	4,559	4,389	4,253	4,142	3,971	3,844	3,747	3,670	3,607	3,412	3,311	3,248	3,119	2,984
18	10,218	7,215	6,028	5,375	4,956	4,663	4,445	4,276	4,141	4,030	3,860	3,734	3,637	3,560	3,498	3,303	3,201	3,139	3,009	2,873
19	10,073	7,093	5,916	5,268	4,853	4,561	4,345	4,177	4,043	3,933	3,763	3,638	3,541	3,464	3,402	3,208	3,106	3,043	2,913	2,776
20	9,944	6,986	5,818	5,174	4,762	4,472	4,257	4,090	3,956	3,847	3,678	3,553	3,457	3,380	3,318	3,123	3,021	2,959	2,828	2,690
21	9,830	6,891	5,730	5,091	4,681	4,393	4,179	4,013	3,880	3,771	3,602	3,478	3,382	3,305	3,243	3,049	2,947	2,884	2,753	2,614
22	9,727	6,806	5,652	5,017	4,609	4,322	4,109	3,944	3,812	3,703	3,535	3,411	3,315	3,239	3,176	2,982	2,880	2,817	2,685	2,546
23	9,635	6,730	5,582	4,950	4,544	4,259	4,047	3,882	3,750	3,642	3,475	3,351	3,255	3,179	3,116	2,922	2,820	2,756	2,624	2,484
24	9,551	6,661	5,519	4,890	4,486	4,202	3,990	3,826	3,695	3,587	3,420	3,296	3,201	3,125	3,062	2,868	2,765	2,702	2,569	2,428
25	9,475	6,598	5,461	4,835	4,433	4,150	3,939	3,776	3,645	3,537	3,370	3,247	3,151	3,075	3,013	2,819	2,716	2,652	2,519	2,377
26	9,406	6,541	5,409	4,785	4,384	4,103	3,893	3,730	3,599	3,492	3,325	3,202	3,107	3,031	2,968	2,774	2,671	2,607	2,473	2,330
27	9,342	6,488	5,361	4,740	4,340	4,059	3,850	3,687	3,557	3,450	3,284	3,161	3,066	2,990	2,928	2,733	2,630	2,565	2,431	2,287
28	9,284	6,440	5,317	4,698	4,300	4,020	3,811	3,649	3,519	3,412	3,246	3,123	3,028	2,952	2,890	2,695	2,592	2,527	2,392	2,247
29	9,230	6,396	5,276	4,659	4,262	3,983	3,775	3,613	3,483	3,377	3,211	3,088	2,993	2,917	2,855	2,660	2,557	2,492	2,357	2,210
30	9,180	6,355	5,239	4,623	4,228	3,949	3,742	3,580	3,450	3,344	3,179	3,056	2,961	2,885	2,823	2,628	2,524	2,459	2,323	2,176
40	8,828	6,066	4,976	4,374	3,986	3,713	3,509	3,350	3,222	3,117	2,953	2,831	2,737	2,661	2,598	2,401	2,296	2,230	2,088	1,932
50	8,626	5,902	4,826	4,232	3,848	3,578	3,376	3,219	3,092	2,988	2,825	2,703	2,609	2,533	2,470	2,272	2,164	2,097	1,951	1,786
60	8,494	5,795	4,729	4,140	3,760	3,492	3,291	3,134	3,008	2,904	2,742	2,620	2,526	2,450	2,387	2,187	2,079	2,010	1,861	1,689
70	8,402	5,720	4,661	4,076	3,698	3,431	3,232	3,075	2,950	2,846	2,684	2,563	2,468	2,392	2,329	2,128	2,019	1,949	1,797	1,618
80	8,334	5,665	4,611	4,029	3,652	3,387	3,188	3,032	2,907	2,803	2,641	2,520	2,425	2,349	2,286	2,084	1,974	1,903	1,748	1,563
90	8,282	5,623	4,573	3,992	3,617	3,352	3,154	2,999	2,874	2,770	2,609	2,487	2,393	2,316	2,253	2,051	1,939	1,868	1,711	1,520
100	8,240	5,589	4,542	3,963	3,589	3,325	3,127	2,972	2,847	2,744	2,583	2,461	2,367	2,290	2,227	2,024	1,912	1,840	1,681	1,485
∞	7,879	5,298	4,279	3,715	3,350	3,091	2,897	2,744	2,621	2,519	2,358	2,237	2,142	2,064	2,000	1,789	1,669	1,590	1,402	1,000

STUDENT (t)-Verteilung:



Für verschiedene Freiheitsgrade ν und für verschiedene Wahrscheinlichkeiten $P(T<t)=1-\alpha$ sind die entsprechenden t-Werte tabelliert (Quelle: J. Schwarze, Grundlagen der Statistik II, 1993, S. 308 - 309)

ν	1 − α									
	0,6000	0,7000	0,8000	0,9000	0,9500	0,9750	0,9900	0,9950	0,9990	0,9995
1	0,3249	0,7265	1,3764	3,0777	6,3137	12,7062	31,8205	63,6568	318,3127	636,5894
2	0,2887	0,6172	1,0607	1,8856	2,9200	4,3027	6,9646	9,9248	22,3273	31,5983
3	0,2767	0,5844	0,9785	1,6377	2,3534	3,1825	4,5407	5,8409	10,2146	12,9238
4	0,2707	0,5686	0,9410	1,5332	2,1318	2,7763	3,7470	4,6041	7,1732	8,6102
5	0,2672	0,5594	0,9195	1,4759	2,0150	2,5706	3,3648	4,0322	5,8934	6,8688
6	0,2648	0,5534	0,9057	1,4398	1,9432	2,4469	3,1426	3,7074	5,2076	5,9588
7	0,2632	0,5491	0,8960	1,4149	1,8946	2,3646	2,9979	3,4995	4,7851	5,4079
8	0,2619	0,5459	0,8889	1,3968	1,8595	2,3060	2,8965	3,3554	4,5007	5,0411
9	0,2610	0,5435	0,8834	1,3830	1,8331	2,2622	2,8214	3,2498	4,2968	4,7808
10	0,2602	0,5415	0,8791	1,3722	1,8125	2,2281	2,7638	3,1693	4,1437	4,5868
11	0,2596	0,5399	0,8755	1,3634	1,7959	2,2010	2,7181	3,1058	4,0247	4,4369
12	0,2590	0,5386	0,8726	1,3562	1,7823	2,1788	2,6810	3,0545	3,9296	4,3178
13	0,2586	0,5375	0,8702	1,3502	1,7709	2,1604	2,6503	3,0123	3,8520	4,2208
14	0,2582	0,5366	0,8681	1,3450	1,7613	2,1448	2,6245	2,9768	3,7874	4,1404
15	0,2579	0,5357	0,8662	1,3406	1,7530	2,1314	2,6025	2,9467	3,7328	4,0727
16	0,2576	0,5350	0,8647	1,3368	1,7459	2,1199	2,5835	2,9208	3,6862	4,0150
17	0,2573	0,5344	0,8633	1,3334	1,7396	2,1098	2,5669	2,8982	3,6458	3,9651
18	0,2571	0,5338	0,8620	1,3304	1,7341	2,1009	2,5524	2,8784	3,6105	3,9216
19	0,2569	0,5333	0,8610	1,3277	1,7291	2,0930	2,5395	2,8609	3,5794	3,8834
20	0,2567	0,5329	0,8600	1,3253	1,7247	2,0860	2,5280	2,8453	3,5518	3,8495
21	0,2566	0,5325	0,8591	1,3232	1,7207	2,0796	2,5176	2,8314	3,5272	3,8193
22	0,2564	0,5321	0,8583	1,3212	1,7171	2,0739	2,5083	2,8188	3,5050	3,7921
23	0,2563	0,5317	0,8575	1,3195	1,7139	2,0687	2,4999	2,8073	3,4850	3,7676
24	0,2562	0,5314	0,8569	1,3178	1,7109	2,0639	2,4922	2,7969	3,4668	3,7454
25	0,2561	0,5312	0,8562	1,3163	1,7081	2,0595	2,4851	2,7874	3,4502	3,7251

ν	$1-\alpha$								
	0,6000	0,7000	0,8000	0,9000	0,9500	0,9750	0,9900	0,9950	0,9990	0,9995
26	0,2560	0,5309	0,8557	1,3150	1,7056	2,0555	2,4786	2,7787	3,4350	3,7066
27	0,2559	0,5306	0,8551	1,3137	1,7033	2,0518	2,4727	2,7707	3,4210	3,6896
28	0,2558	0,5304	0,8546	1,3125	1,7011	2,0484	2,4671	2,7633	3,4082	3,6739
29	0,2557	0,5302	0,8542	1,3114	1,6991	2,0452	2,4620	2,7564	3,3962	3,6594
30	0,2556	0,5300	0,8538	1,3104	1,6973	2,0423	2,4573	2,7500	3,3852	3,6459
31	0,2555	0,5298	0,8534	1,3095	1,6955	2,0395	2,4528	2,7440	3,3749	3,6334
32	0,2555	0,5297	0,8530	1,3086	1,6939	2,0369	2,4487	2,7385	3,3653	3,6218
33	0,2554	0,5295	0,8526	1,3077	1,6924	2,0345	2,4448	2,7333	3,3563	3,6109
34	0,2553	0,5294	0,8523	1,3070	1,6909	2,0322	2,4411	2,7284	3,3479	3,6007
35	0,2553	0,5292	0,8520	1,3062	1,6896	2,0301	2,4377	2,7238	3,3400	3,5911
36	0,2552	0,5291	0,8517	1,3055	1,6883	2,0281	2,4345	2,7195	3,3326	3,5821
37	0,2552	0,5289	0,8514	1,3049	1,6871	2,0262	2,4314	2,7154	3,3256	3,5737
38	0,2551	0,5288	0,8512	1,3042	1,6860	2,0244	2,4286	2,7116	3,3190	3,5657
39	0,2551	0,5287	0,8509	1,3036	1,6849	2,0227	2,4258	2,7079	3,3128	3,5581
40	0,2550	0,5286	0,8507	1,3031	1,6839	2,0211	2,4233	2,7045	3,3069	3,5509
41	0,2550	0,5285	0,8505	1,3025	1,6829	2,0195	2,4208	2,7012	3,3013	3,5442
42	0,2550	0,5284	0,8503	1,3020	1,6820	2,0181	2,4185	2,6981	3,2960	3,5377
43	0,2549	0,5283	0,8501	1,3016	1,6811	2,0167	2,4163	2,6951	3,2909	3,5316
44	0,2549	0,5282	0,8499	1,3011	1,6802	2,0154	2,4141	2,6923	3,2861	3,5258
45	0,2549	0,5281	0,8497	1,3006	1,6794	2,0141	2,4121	2,6896	3,2815	3,5202
46	0,2548	0,5281	0,8495	1,3002	1,6787	2,0129	2,4102	2,6870	3,2771	3,5149
47	0,2548	0,5280	0,8493	1,2998	1,6779	2,0117	2,4083	2,6846	3,2729	3,5099
48	0,2548	0,5279	0,8492	1,2994	1,6772	2,0106	2,4066	2,6822	3,2689	3,5051
49	0,2547	0,5278	0,8490	1,2991	1,6765	2,0096	2,4049	2,6800	3,2651	3,5004
50	0,2547	0,5278	0,8489	1,2987	1,6759	2,0086	2,4033	2,6778	3,2614	3,4960
60	0,2545	0,5272	0,8477	1,2958	1,6706	2,0003	2,3901	2,6603	3,2317	3,4602
70	0,2543	0,5268	0,8468	1,2938	1,6669	1,9944	2,3808	2,6479	3,2108	3,4350
80	0,2542	0,5265	0,8461	1,2922	1,6641	1,9901	2,3739	2,6387	3,1953	3,4163
90	0,2541	0,5263	0,8456	1,2910	1,6620	1,9867	2,3685	2,6316	3,1833	3,4019
100	0,2540	0,5261	0,8452	1,2901	1,6602	1,9840	2,3642	2,6259	3,1737	3,3905

VERZEICHNIS WICHTIGER SYMBOLE

Übersicht über wichtige Symbole und das griechische Alphabet

Symbol	Bedeutung
$\Phi(x)$	Verteilungsfunktion der Standardnormalverteilung N(0,1)
$\varphi(x)$	Dichtefunktion der Standardnormalverteilung N(0,1)
$\sim$	„verteilt nach“, z.B. X ~ N(0,1) = „X ist standardnormalvetrteilt“
$\mu, auch\ \mu_X\ oder\ E(X)$	Erwartungswert der Zufallsvariablen X
$\sigma^2, auch\ \sigma_X^2\ oder Var(X)$	Varianz der Zufallsvariablen X
$Cov(X,Y)$	Covarianz der Zufallsvariablen X und Y
$N(\mu,\sigma^2)$	Normalverteilung mit Mittelwert μ und Varianz σ^2
$N(0,1)$	Standardnormalverteilung mit Mittelwert 0 und Varianz 1
t_υ	(Student-) t-Verteilung mit υ Freiheitsgraden
χ_υ^2	Chiquadrat-Verteilung mit υ Freiheitsgraden
$F_{\upsilon_1,\upsilon_2}$	F-Verteilung mit υ_1 Zähler – und υ_2 Nennerfreiheitsgraden
$B(n,p)$	Binomialverteilung mit den Parametern n (Anzahl Versuche) und p (Wahrscheinlichkeit „Treffer“)
$G, auch\ \Omega$	Ereignisraum (Menge aller möglichen Ergebnisse eines Zufallsexperiments)
$P(A)$	Wahrscheinlichkeit für den Eintritt des Erignisses A
$P(A \mid B)$	Bedingte Wahrscheinlichkeit für den Eintritt des Ereignisses A unter der Bedingung des Eintretens des Ereignisses B
$F(x)$	Verteilungsfunktion der Zufallsvariablen X
$f(x)$	Dichtefunktion der Zufallsvariablen X
$\rho_{x,y}$	Theoretische Korrelation der Zufallsvariablen X und Y
$\bar{x}$	aritmetisches Mittel (aus der Stichprobe)
s^2, s_x^2	empirische Varianz (aus der Stichprobe)
$s_{x,y}$	empirische Covarianz (aus der Stichprobe)
$r_{x,y}$	empirische Korrelation (aus der Stichprobe)
R^2	Bestimmtheitsmaß (der Regression)
ML	Maximum Likelihood
$F(x,y)$	gemeinsame Verteilungsfunktion der Zufallsvariablen X und Y
$f(x,y)$	gemeinsame Dichtefunktion der der Zufallsvariablen X und Y
$\ln x$	Natürlicher Logarithmus (zur Basis e = 2,7182818…)

$\log x$	Dekadische Logarithmus (zur Basis 10)
$e^x, \exp(x)$	e-Funktion , Exponentialfunktion zur Basis e
$n!$	„n Fakultät“: n! = 1*2*…*(n-1)*n
$\binom{n}{k}$	Binomialkoeffizient („n über k“) = $\frac{n!}{k!(n-k)!}$
$\lim_{x\to\infty}$	Grenzwert (Limes) wenn x gegen unendlich geht
$a_n \underset{n\to\infty}{\to} a$	„a konvergiert gegen a, wenn n gegen unendlich läuft“
min	Minimum, kleinster Wert einer Menge
max	Maximum, größter Wert einer Menge
$\lvert x\rvert$	Absolutwert, Betrag von x
$\sum_i x_i, \sum_i x_i$	Summe von x über alle i = $x_1 + x_2 + \ldots + x_{n-1} + x_n$
$\sum_i \sum_j x_{ij}, \sum_i \sum_j x_{ij}$	Doppelsumme über x und y = $x_{11} + x_{12} + \ldots + x_{1m} + x_{21} + x_{22} + \ldots + x_{2m} + \ldots + x_{n1} + x_{n2} + \ldots + x_{nm}$
$\prod_i x_i$	Produkt von x über alle i = $x_1 * x_2 * \ldots * x_n$

Rechenzeichen

$=$	„gleich“
$\approx$	„ungefähr gleich“, „rund“
$\hat{=}$	„entspricht“
$\leq$	„kleiner gleich“
$\geq$	„größer gleich“
$\in$	Element von; z.B. $3 \in \{1;2;3;4\}$
$\notin$	ist nicht Element von; z.B. $3 \notin \{1;2;4\}$
$\subset$	„Teilmenge von“ $\{1;3\} \subset \{1;2;3;4\}$
$\varnothing$	„unmögliches Ereignis“, „leere Menge“
$\cup$	„vereinigt mit“
$\cap$	„geschnitten mit“

Griechisches Alphabet

$\mathrm{A}\alpha$	Alpha
$\mathrm{B}\beta$	Beta
$\Gamma\gamma$	Gamma
$\Delta\delta$	Delta
$\mathrm{E}\varepsilon$	Epsilon
$\mathrm{Z}\zeta$	Zeta
$\mathrm{H}\eta$	Eta
$\Theta\theta$	Theta
$\mathrm{I}\iota$	Jota
$\mathrm{K}\kappa$	Kappa
$\Lambda\lambda$	Lambda

$\mathrm{M}\mu$	My
$\mathrm{N}\nu$	Ny
$\Xi\xi$	Xi
$\mathrm{O}o$	Omikron
$\Pi\pi$	Pi
$\mathrm{P}\rho$	Rho
$\Sigma\sigma$	Sigma
$\mathrm{T}\tau$	Tau
$\Upsilon\upsilon$	Ypsilon
$\Phi\phi\varphi$	Phi
$\mathrm{X}\chi$	Chi
$\Psi\psi$	Psi
$\Omega\omega$	Omega