

Empirische Wirtschaftsforschung

Logit- und Probitmodelle

Ulrich Fritsche

Universität Hamburg

Email: Ulrich.Fritsche@wiso.uni-hamburg.de

Idee

Logit- und Probitschätzer als Modelle mit begrenzten abhängigen Variablen, multivariate Antwortmodelle.

Im binären Fall lässt sich schreiben:

$$y_i = \begin{cases} 1 & : \text{bestimmtes Ereignis tritt ein, Wahl wird getroffen} \\ 0 & : \text{für das Gegenteil} \end{cases}$$

Beispiele:

- ▶ Studium: Ja/Nein?
- ▶ Wohnung: Mieten/Kaufen?

Frage: Was bestimmt die Wahrscheinlichkeit p , dass eine bestimmte Entscheidung getroffen wird, bzw. ein bestimmtes Ereignis passiert?

p ist nicht beobachtbar, aber y , d.h., das Ergebnis der Entscheidung

Die Wahrscheinlichkeit p hängt von einer Reihe erklärender Variablen ab, wobei es entscheidend auf die spezifische Form der Funktion F ankommt.

In allgemeiner Form¹

$$p_i \equiv P[y_i = 1] = F(\mathbf{X}\beta) \quad (1)$$

¹Erinnerung: Das OLS-Modell in Matrixschreibweise: $\mathbf{y} = \mathbf{X}\beta + \mathbf{u}$. 

Nimmt man für F eine einfache lineare Funktion an, so kann man mit OLS schätzen:

$$p_i = \mathbf{X}\beta + \mathbf{u} \quad (2)$$

Allerdings weist dieser Ansatz eine Reihe von Problem auf, die sich anhand eines einfachen Beispiels darstellen lassen:

Beispiel

Schätze in GRETL

$$GRAD = c + ASVABC + u \quad (3)$$

mit *GRAD* als binäre Variable: 1 für Diplom bestanden, 0 für nicht bestanden; und *ASVABC* als Ergebnis eines Intelligenztests.

Tabelle: Lineares Wahrscheinlichkeitsmodell

Dependent Variable: GRAD

Method: Least Squares

Sample: 1 540

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.406327	0.057584	7.056208	0.000
ASVABC	0.010187	0.001096	9.292613	0.000

Das lineare Wahrscheinlichkeitsmodell - Problem 1

Wahrscheinlichkeiten größer 1

Interpretation des Schätzergebnisses:

- ▶ Konstante: Wenn Testergebnis 0, Wahrscheinlichkeit des Abschlusses 40%
- ▶ β_1 : Erhöhung des Testergebnisses um 1 Standardabweichung erhöht Wahrscheinlichkeit des Abschlusses um 10%.²

Dies bedeutet, dass für alle Teilnehmer mit Testergebnis besser als 60 die Wahrscheinlichkeit eines Abschlusses größer als 1 ist !

²Der Test ist skaliert mit Standardabweichung 10.

Das lineare Wahrscheinlichkeitsmodell - Problem 2a

Heteroskedastische Residuen

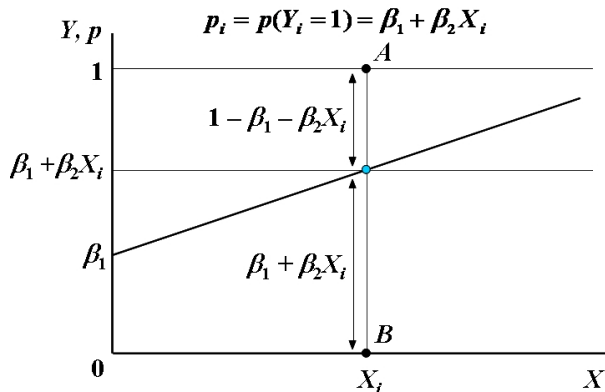


Abbildung: Heteroskedastische Residuen

Problem 2a (cont'd)

Heteroskedastische Residuen

Varianz des Störterms :

$$\text{Var}(\mathbf{u}) = E(\mathbf{u}\mathbf{u}') = (1 - \mathbf{X}\mathbf{b})(\mathbf{X}\mathbf{b}) \quad (4)$$

verglichen mit der Formulierung für Standard-OLS³

$$E(\mathbf{u}\mathbf{u}') = \hat{\sigma}^2 \mathbf{I} \quad (5)$$

d.h., die Varianz der Störterme im linearen Wahrscheinlichkeitsmodell verändert sich mit \mathbf{X} ,
⇒ Heteroskedastie.

³Siehe ersten Teil der Vorlesung.

Problem 2b

Residuen nicht normalverteilt

Die Residuen sind um zwei Punkte herum zentriert:

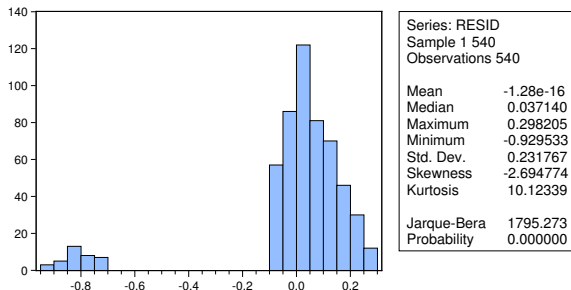


Abbildung: Nichtnormalverteilte Residuen

Jarque-Bera Test

Vergleich der Kurtosis (Wölbung) und Skewness (Schiefe), basierend auf LM Test.⁴

Die Test-Statistik ergibt sich als:

$$LM = N \left[\frac{1}{6} \left(\frac{1}{T} \sum_{t=i}^N \frac{\hat{u}_i^3}{\hat{\sigma}^3} \right)^2 + \frac{1}{24} \left(\frac{1}{N} \sum_{t=i}^N \frac{\hat{u}_i^4}{\hat{\sigma}^4} - 3 \right)^2 \right] \quad (6)$$

Unter

$$H_0 : \mathbf{u} \sim N$$

⁴Siehe Verbeek (2008), S.195

Lösungen

- ▶ Annahme anderer spezifischer Funktionen für F :
Reskalierung der Wahrscheinlichkeitsverteilung, Behebung von Problem 1.
Definiere $Z = \mathbf{X}\beta$, mit $P = F(Z)$.
- ▶ Verwendung Maximum-Likelihood um nichtlineare Funktionen schätzen zu können

Das Logit Modell

Für $Z = \mathbf{X}\beta$ und $P = F(Z)$, verwende eine logistische Funktion für F

$$F(Z) = \frac{1}{1 + e^{-Z}} = \frac{1}{1 + e^{-\mathbf{X}\beta}} \quad (7)$$

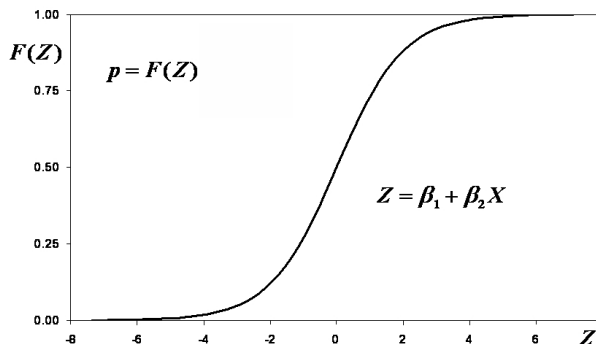
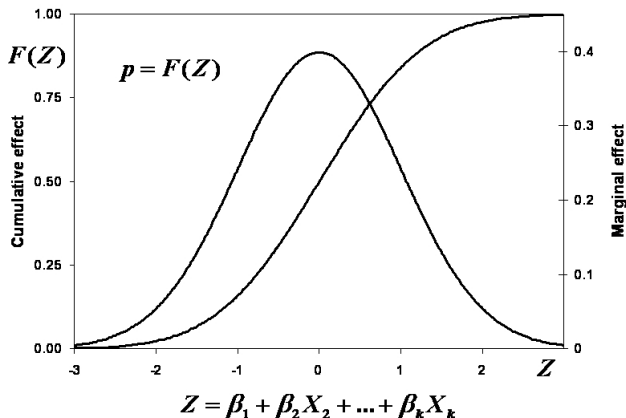


Abbildung: Logistische Funktion

Das Probit Modell

Unter Annahme der kumulativen Standardnormalverteilung für F erhält man

$$F(Z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^Z e^{-\frac{1}{2}Z^2} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\mathbf{x}\beta} e^{-\frac{1}{2}\mathbf{x}\beta} \quad (8)$$



Interpretation

Um die Schätzer des Logit- und Probit Modells interpretieren zu können, berechnet man jeweils die ▶ 1. Ableitung:

$$f(Z)\beta = \frac{dp}{dZ}\beta = \frac{e^{-\mathbf{x}\beta}}{(1 + e^{-\mathbf{x}\beta})^2}\beta \quad (9)$$

und

$$f(Z)\beta = \frac{dp}{dZ}\beta = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(\mathbf{x}\beta)^2}\beta \quad (10)$$

Wichtig: Die Effekte hängen vom Wert von \mathbf{X} ab. Meist wird hierfür der Durchschnitt der Beobachtungsreihe verwendet.

Interpretation des Schätzers als Regime-Switch

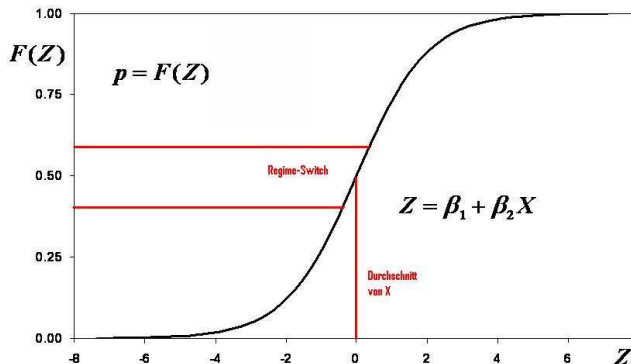


Abbildung: Logistische Funktion

Schätzung mit Maximum Likelihood

Da die Gleichungen (9) und (10) nichtlinear sind, wird der ML-Schätzer verwendet:

Die Log-Likelihood Funktion ist gegeben als⁵

$$\log L(\beta) = \sum_{i=1}^N \log F(\mathbf{X}_i\beta) + \sum_{i=1}^N (1 - y_i) \log(1 - F(\mathbf{X}_i\beta)) \quad (11)$$

Die 1. Ableitung führt auf den ML-Schätzer:

$$\frac{\partial \log L(\beta)}{\partial \beta} = \sum_{i=1}^N \left[\frac{y_i - F(\mathbf{X}_i\beta)}{F(\mathbf{X}_i\beta)(1 - F(\mathbf{X}_i\beta))} f(\mathbf{X}_i\beta) \right] \mathbf{X}_i = 0 \quad (12)$$

mit $f = F'$ als Ableitung der unterstellten Funktion.

⁵Siehe Verbeek (2008), S.203.

Für das Logit Modell ergibt sich: Herleitung

$$\frac{\partial \log L(\beta)}{\partial \beta} = \sum_{i=1}^N \left[y - \frac{e^{\mathbf{x}\beta}}{1 + e^{\mathbf{x}\beta}} \right] \mathbf{x} = 0 \quad (13)$$

Das heißt, wir erhalten den Schätzer für $p(y = 1)$ bei gegebenem \mathbf{X} als:

$$\hat{p} = \frac{e^{\mathbf{x}\hat{\beta}}}{1 + e^{\mathbf{x}\hat{\beta}}} \quad (14)$$

Für die Bewertung der Güte der Schätzung verwendet man das **McFadden R^2** :

$$McFaddenR^2 = 1 - \frac{\log L_1}{\log L_0} \quad (15)$$

mit L_1 als Log-Likelihood des geschätzten Modells, und L_0 als Log-Likelihood eines Modells mit einer Konstanten.

Beispiel: Schätzung von Logit und Probit

Schätzung von Gleichung (3) in Gretl mit Logit- und Probit-Modell:

$$GRAD = c + ASVABC + u \quad (3)$$

Dies führt auf:

Variable	Logit	Probit
C	-5.829	-3.03215
ASVABC	0.186413	0.09891

Dies sind jedoch noch nicht die Koeffizienten! Wir müssen den marginalen Effekt berechnen, :

$$\frac{\partial p}{\partial AVSABC} = \frac{\partial p}{\partial Z} \frac{\partial Z}{\partial ASVABC} = f(Z)\beta \quad (16)$$

Aus den Gleichungen (9) und (10):

$$f(Z)\beta = \frac{e^{-\mathbf{x}\beta}}{(1 + e^{-\mathbf{x}\beta})^2}\beta \quad (9)$$

und

$$f(Z)\beta = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\mathbf{x}\beta)^2}\beta \quad (10)$$

Für $\mathbf{X}\beta$ verwenden wir den Durchschnitt, d.h. im Logit-Fall:

$$\begin{aligned}\mathbf{X}\beta &= \bar{\mathbf{x}}_1 \times \hat{\beta}_1 + \bar{\mathbf{x}}_2 \times \hat{\beta}_2 \\ &= 1 \times -5.829002 + 51.73 \times 0.186413 \\ &= 3.81414249\end{aligned}\tag{17}$$

Analog erhält man für das Probit-Modell:

$$\begin{aligned}\mathbf{X}\beta &= \bar{\mathbf{x}}_1 \times \hat{\beta}_1 + \bar{\mathbf{x}}_2 \times \hat{\beta}_2 \\ &= 1 \times -3.032147 + 51.73 \times 0.09891 \\ &= 2.0844673\end{aligned}\tag{18}$$

Damit erhält man $e^{-\mathbf{X}\beta} = 0.02206$ (Logit) und $e^{-\frac{1}{2}(\mathbf{X}\beta)^2} = 0.11389$ (Probit).

Einsetzen in (9) ergibt den Logit-Schätzer für *ASVABC* als

$$\begin{aligned} f(\mathbf{X}\beta)\beta_1 &= \frac{e^{-\mathbf{X}\hat{\beta}}}{(1 + e^{-\mathbf{X}\hat{\beta}})^2} \hat{\beta}_1 \\ &= \frac{0.02206}{(1 + 0.02206)^2} \times 0.186413 \\ &= 0.003936 \end{aligned} \tag{19}$$

sowie in (10) den Probit-Schätzer für *ASVABC*:

$$\begin{aligned} f(\mathbf{X}\beta)\beta_1 &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\mathbf{X}\hat{\beta})^2} \hat{\beta}_1 \\ &= 0.3989 \times 0.11398 \times 0.09891 \\ &= 0.004493 \end{aligned} \tag{20}$$

Erweiterungen

- ▶ Multivariate Antwortmodelle
- ▶ Modelle mit stetigen Variablen
- ▶ z-Statistik als asymptotische t-Statistik
- ▶ Numerische Verfahren zur Berechnung der Schätzer
- ▶ Tobit Modelle: Abhängige Variable ist begrenzt, "Selection Bias"

Anwendungsbeispiel

Stadtplanung: Abschätzung der Rentabilität des Baus einer U-Bahnstation

Domencich / McFadden (1976): *Urban Travel Demand: A Behavioral Analysis*

Schätze ein Logit-Modell für $P[Auto] = 1$

$$P[Auto] = -3.82 + 0.158TW - 0.382(AIV - TSS) - 2.56(AC - F) + 4.94A/W - 2.91R - 2.36Z \quad (21)$$

mit:

TW = transit time (in minutes)

AIV = auto in-vehicle time (in minutes)

TSS = transit station-to-station (in minutes)

AC = auto parking charges (in US-Dollar)

F = transit fare (in US-Dollar)

A/W = autos per worker in household

R = race (0: white, 1: non-white)

Z = profession (0: blue-collar worker, 1: white-collar worker)

Zusammenfassung

- ▶ Binäre Antwortmodelle zur Beschreibung von begrenzten Variablen
- ▶ Probleme des OLS-Ansatzes
- ▶ Logit und Probit Modelle als einfachste Schätzansätze
- ▶ Anwendungsgebiet von Maximum Likelihood
- ▶ **Literatur:**
 - Dougherty (2007), Kapitel 11
 - Verbeek (2008), Kapitel 7
 - Winker (2008), Kapitel 9.2

Hintergrundfolien

Quotientenregel:

$$F(x) = \frac{u(x)}{v(x)}$$

$$f(x) = \frac{u'v - uv'}{v^2}$$

also:

$$\begin{aligned} f(Z)\beta &= \left(\frac{0 \times (1 + e^{-\mathbf{x}\beta}) - 1 \times (-e^{-\mathbf{x}\beta})}{(1 + e^{-\mathbf{x}\beta})^2} \right) \beta \\ &= \frac{e^{-\mathbf{x}\beta}}{(1 + e^{-\mathbf{x}\beta})^2} \beta \end{aligned}$$

Einsetzen führt auf:

$$\begin{aligned}
 \frac{\partial \log L(\beta)}{\partial \beta} &= \left[\frac{\mathbf{y} - \left(\frac{1}{1+e^{-\mathbf{x}\beta}} \right) \frac{e^{-\mathbf{x}\beta}}{\frac{1}{1+e^{-\mathbf{x}\beta}} - \frac{1}{(1+e^{-\mathbf{x}\beta})^2}}}{(1+e^{-\mathbf{x}\beta})^2} \right] \mathbf{X} \\
 &= \left[\frac{\left(\mathbf{y} - \left(\frac{1}{1+e^{-\mathbf{x}\beta}} \right) \right) e^{-\mathbf{x}\beta}}{\frac{(1+e^{-\mathbf{x}\beta})^2}{1+e^{-\mathbf{x}\beta}} - \frac{(1+e^{-\mathbf{x}\beta})^2}{(1+e^{-\mathbf{x}\beta})^2}} \right] \mathbf{X} \\
 &= \left[\frac{\left(\mathbf{y} - \left(\frac{1}{1+e^{-\mathbf{x}\beta}} \right) \right) e^{-\mathbf{x}\beta}}{1+e^{-\mathbf{x}\beta} - 1} \right] \mathbf{X} \\
 &= \left[\mathbf{y} - \frac{1}{1+e^{-\mathbf{x}\beta}} \right] \mathbf{X}
 \end{aligned}$$