

Logit-Modelle für ordinale und multinomiale Variablen

Statistik II

Ordinale abhängige Variablen
Kategoriale abhängige Variablen
Zusammenfassung

Warum binäre Logit-Modelle?
Ordinale abhängige Variablen
Latent Variable Interpretation
Kategoriale abhängige Variablen
Zusammenfassung

Binäre Logit-Modelle

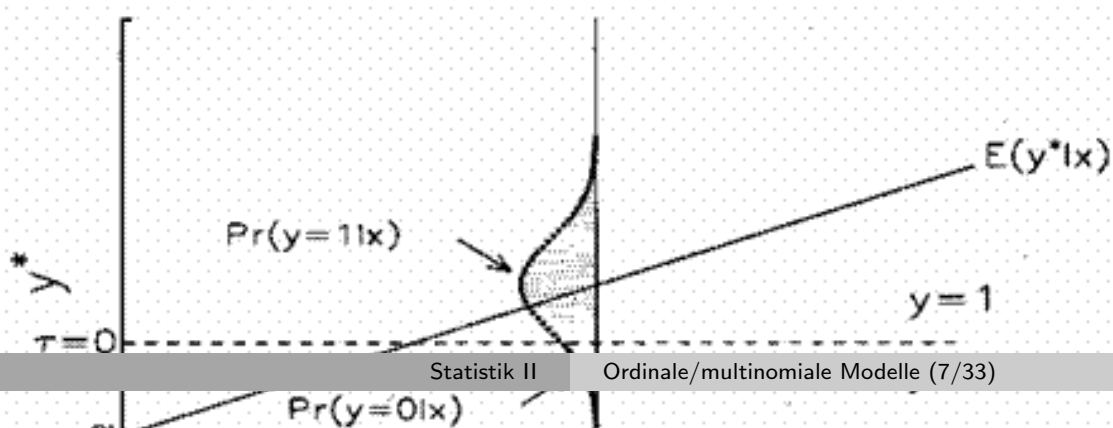
- ▶ Viele interessante abhängige Variablen binär
 - ▶ Interpretation als Wahrscheinlichkeiten
 - ▶ Wahrscheinlichkeiten < 0 oder > 1
 - ▶ Heteroskedastizität
 - ▶ Linearer Zusammenhang unplausibel
- ▶ Logit-Transformation
 - ▶ Wahrscheinlichkeit durch Gegenwahrscheinlichkeit teilen:
„odds“ mit positivem Wertebereich
 - ▶ Natürlicher Logarithmus von odds: „logits“ mit komplettem Wertebereich
 - ▶ Nicht-lineare Transformation der abhängigen Variablen
- ▶ Modell linear in den Logits
- ▶ Transformation umkehren – erwartete Wahrscheinlichkeiten, nicht-linearer Zusammenhang mit x_1, x_2, \dots

Was sind latente Variable?

- ▶ Sehr viele Variablen „latent“ – nicht direkt beobachtbar
- ▶ Alle Einstellungsvariablen (Parteibindung, Ideologie, etc.)
- ▶ Beobachtbar: Auswirkungen/Konsequenzen/Indikatoren (i. d. R. (politisches) Verhalten)
- ▶ Beispiel CDU-Wahl
 - ▶ Alle Befragten haben mehr oder minder große „Tendenz“ CDU zu wählen
 - ▶ Bei Befragten mit großer Tendenz wird CDU-Wahl beobachtet (1)
 - ▶ Bei Befragten mit geringer Tendenz wird keine CDU-Wahl beobachtet (0)
- ▶ Ok?
- ▶ Wie groß muß Tendenz sein? Zufällige Einflüsse?

Zufällige Einflüsse, y^* und y

- ▶ y^* hängt ab von $\beta_0 + \beta_1 x_1$ (systematische Einflüsse) und ϵ (zufällige Einflüsse)
- ▶ ϵ symmetrisch verteilt mit Mittelwert von 0 (Standardannahme)
- ▶ Wenn $y^* + \epsilon > 0$: $y = 1$
- ▶ Wenn $y^* + \epsilon \leq 0$: $y = 0$
- ▶ Wahrscheinlichkeit von $y = 1 = \text{CDU-Wahl?}$

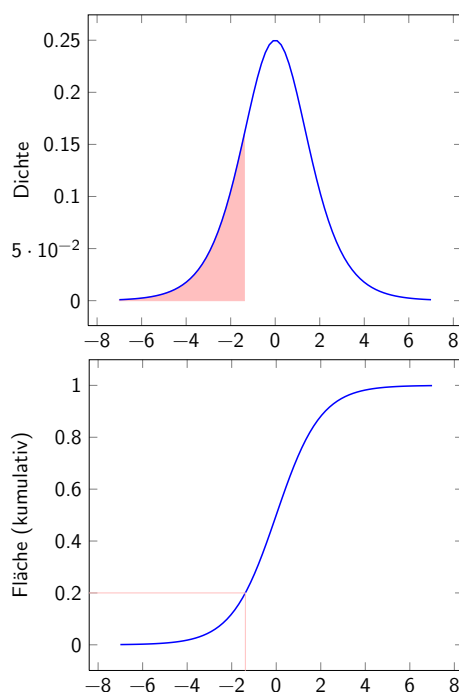


Ordinale abhängige Variablen
Kategoriale abhängige Variablen
Zusammenfassung

Latent Variable Interpretation
Ordered Logit Model

Standardlogistische Verteilung von y^*

Standardlogistische Verteilung

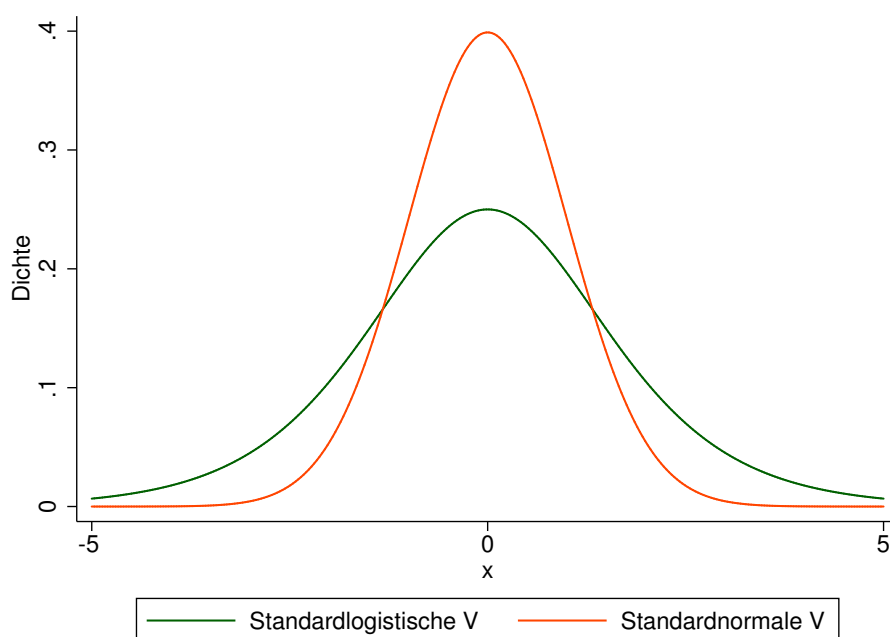


Wahrscheinlichkeitsdichte:

$$f(x) = \frac{\exp(x)}{(1+\exp x)^2}$$

Integral dazu: $f(x) = \frac{\exp(x)}{(1+\exp(x))}$

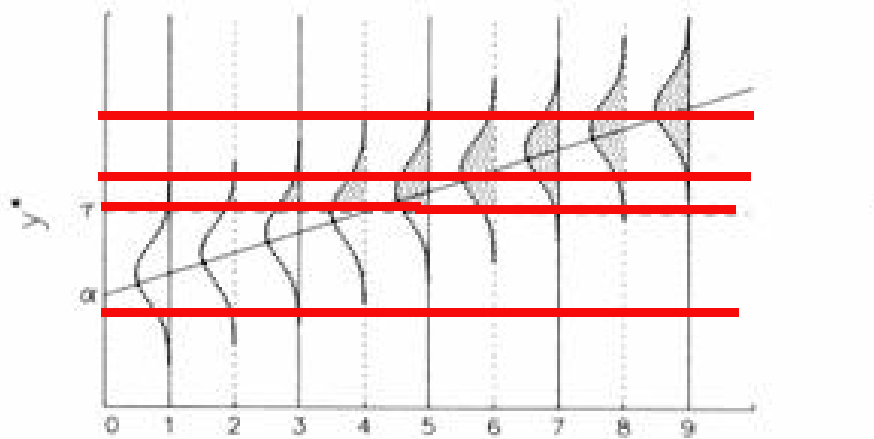
Standardlogistische vs. Standardnormalverteilung



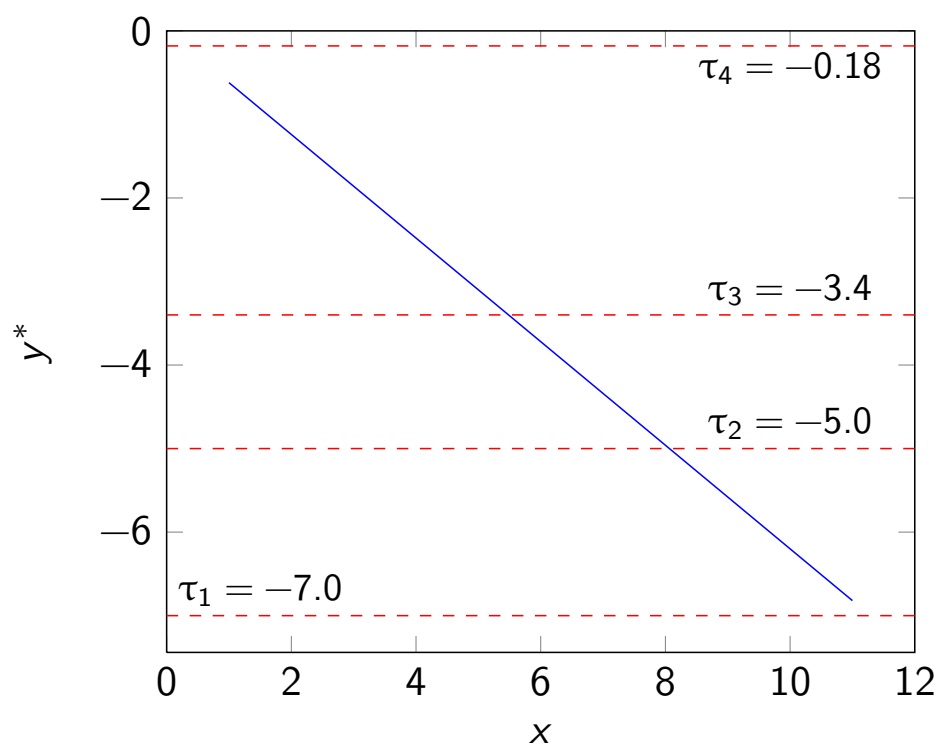
Wieso Schwellenwerte?

- ▶ Für binäres Modell ein Schwellenwert
 $\tau = 0; y = 1$ wenn $y^* > \tau$
- ▶ Für ordinale Modell mehrere Schwellenwerte, da mehr Ausprägungen von y
- ▶ Z. B. vier Schwellenwerte für Variable mit fünf Ausprägungen
- ▶ Wahrscheinlichkeit einer Antwort entspricht Wahrscheinlichkeit, daß y^* zwischen zugehörigen Schwellenwerten liegt
- ▶ Keine erwarteten Werte außerhalb der ordinalen Skala, kein Problem mit Heteroskedastizität
- ▶ Schätzung wieder mit Maximum Likelihood, Annahme über (konditionale) Varianz von y^* bzw. Varianz von ϵ wieder standardlogistische Verteilung

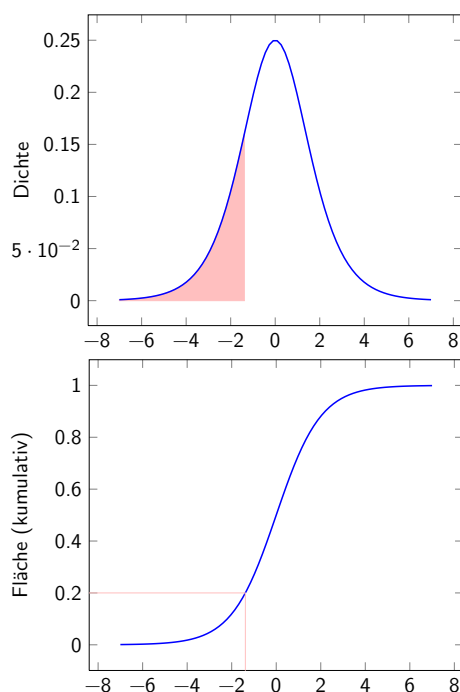
Ordered Logit mit fünf Ausprägungen



Fünf Kategorien, vier Schwellenwerte



Standardlogistische Verteilung



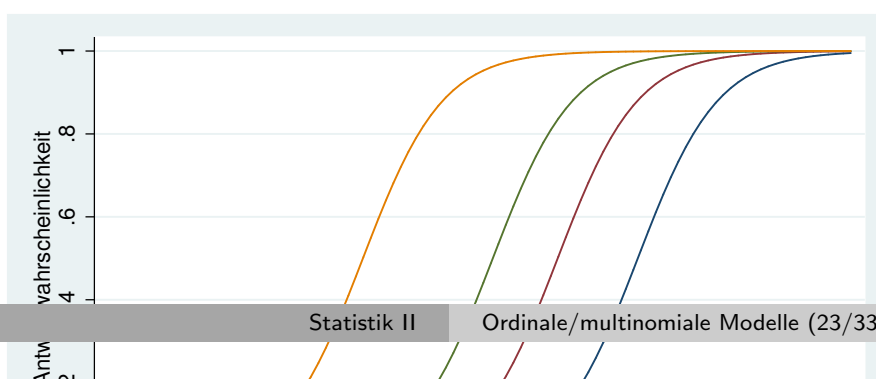
Wahrscheinlichkeitsdichte:

$$f(x) = \frac{\exp(x)}{(1 + \exp(x))^2}$$

Integral dazu: $F(x) = \frac{\exp(x)}{(1 + \exp(x))}$

„Parallele Regression“

- ▶ Fünf Ausprägungen, aber nur ein Effekt für LRS
- ▶ Abstände zwischen Ausprägungen nicht identisch (Schwellenwerte unterschiedlich weit voneinander entfernt)
- ▶ Aber identische Wirkung der LRS auf alle Kontraste (1 vs. 2,3,4,5; 1,2 vs. 3,4,5 ...)
- ▶ Parallel Regression Assumption
- ▶ Statistischer Test möglich (hier keine signifikanten Hinweise auf Verletzung)
- ▶ Bei Bedarf komplexere Modelle möglich



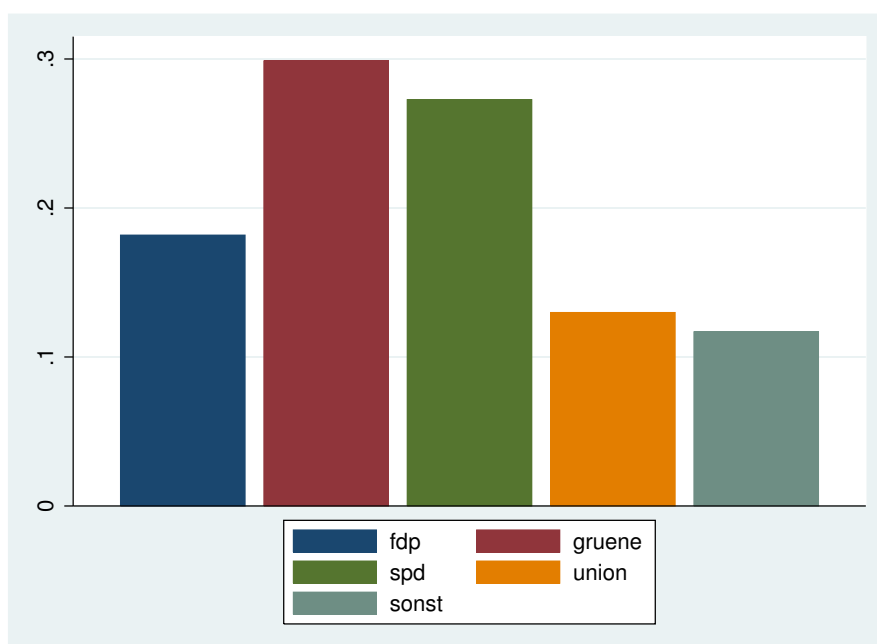
Warum noch ein Modell???

- ▶ Dichotome/binäre bzw. ordinale abhängige Variablen wichtig
- ▶ Ordinales Modell: unterstellt, daß Kategorien in Reihenfolge zusammengefaßt werden können
- ▶ Was ist mit gänzlich ungeordneten Kategorien (z. B. Wahlabsicht)?
- ▶ Weiteres, flexibles Modell nötig
- ▶ Multinomiales Modell flexibel, aber unübersichtlich

Was drücken die Koeffizienten aus?

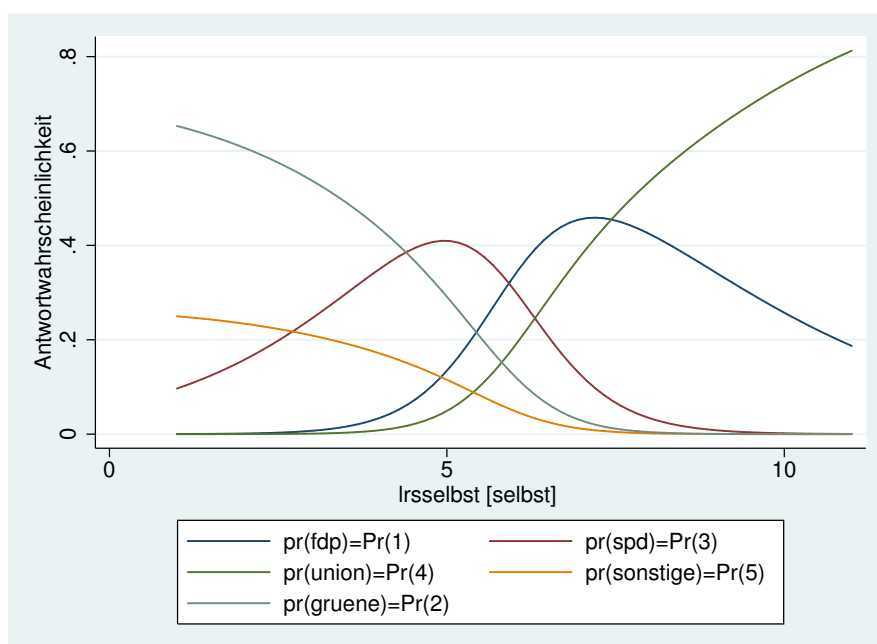
- ▶ Logit von Alternative j vs. Referenzkategorie b (Odds)
- ▶ Bzw. $\ln(\Pr(j)/\Pr(b))$
- ▶ Daraus folgt, das $\beta_b = 0$
 - ▶ Da $\frac{\Pr(b)}{\Pr(b)} = 1$ und
 - ▶ $\ln(1) = 0$
- ▶ Koeffizient für Referenzkategorie muß nicht geschätzt werden, sondern ergibt sich aus Spezifikation des Modells

Beispiel Wahlabsicht und LRS



- ▶ Annahme: Wahlabsicht wird von Ideologie (LRS) beeinflusst
- ▶ Ganz Linke sollten vor allem ..Andere“ und ..Grüne“ wählen
- ▶ Ganz Rechte vermutlich FDP

LRS und erwartetes Wahlverhalten



Zusammenfassung

- ▶ Binäre, ordinale und vor allem kategoriale (polytome) abhängige Variablen:
- ▶ Häufige und gravierende Verletzung der Annahmen des linearen Modells
- ▶ Seit ca. Mitte der 1970er Jahre Modelle für kategoriale Daten in der (amerikanischen) Politikwissenschaft, heute Standard
- ▶ Mathematik hinter Modellen etwas komplex (vor allem für multinomiale Daten)
- ▶ Interpretation trickreich wegen Nicht-Linearitäten und Vielzahl von Koeffizienten
- ▶ Wichtig: Verständnis für Verfahren, (graphische) Interpretation, wann welches Verfahren anwenden