Logit-Modelle für ordinale und multinomiale Variablen

Statistik II

Zum Nachlesen

- Für letzte Stunde: Long/Freese Kapitel 4.1, 4.2 und 4.6 (im Reader unter zusätzliche Dokumente)
- Für diese Woche: Long/Freese Kapitel 5.1, 5.2, 5.6, 6.1, 6.2, 6.6 (im Reader)

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 $x_1, x_2, \cdots$

Was weiter?

- Binäres Logit-Modell funktioniert für dichotome (0/1) abhängige Variablen
- Viele interessante Variablen sehen anders aus

Was weiter?

- Binäres Logit-Modell funktioniert für dichotome (0/1) abhängige Variablen
- Viele interessante Variablen sehen anders aus
  - Was ist mit ordinalen abhängigen Variablen? (lehne stark ab, lehne ab, stimme zu, stimme stark zu)
  - Was ist mit ungeordneten abhängigen Variablen? (CDU, SPD, Grüne, FDP, Linkspartei, Nichtwahl, andere Partei)

Was weiter?

- Binäres Logit-Modell funktioniert für dichotome (0/1) abhängige Variablen
- Viele interessante Variablen sehen anders aus
  - Was ist mit ordinalen abhängigen Variablen? (lehne stark ab, lehne ab, stimme zu, stimme stark zu)
  - Was ist mit ungeordneten abhängigen Variablen? (CDU, SPD, Grüne, FDP, Linkspartei, Nichtwahl, andere Partei)
- Spezielle Modelle, die mit diesen Variablen umgehen können
- Ordered Logit, Multinomial Logit
- Nicht kompliziert, nur etwas komplex

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

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)

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?

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?

Latente Variable und zufällige Einflüsse

- Herleitung des Logit-Modells

Latente Variable und zufällige Einflüsse

- Herleitung des Logit-Modells
- Beobachtete Variable = CDU-Wahl = y = 1 oder 0

Latente Variable und zufällige Einflüsse

- Herleitung des Logit-Modells
- Beobachtete Variable = CDU-Wahl = $y = 1$ oder 0
- Latente Variable = „CDU-Tendenz“ = $y^* =$ Logit

Latente Variable und zufällige Einflüsse

- Herleitung des Logit-Modells
- Beobachtete Variable = CDU-Wahl = $y = 1$ oder 0
- Latente Variable = „CDU-Tendenz“ = $y^* =$ Logit
- Lineares Modell für latente Variable: $y^* = \beta_0 + \beta_1$

Latente Variable und zufällige Einflüsse

- Herleitung des Logit-Modells
- Beobachtete Variable = CDU-Wahl = $y = 1$ oder 0
- Latente Variable = „CDU-Tendenz“ = y^* = Logit
- Lineares Modell für latente Variable: $y^* = \beta_0 + \beta_1$
- Schwellenwert $\tau = 0$
  - Wenn $y^* > \tau: \quad y = 1$
  - Wenn $y^* \leqslant \tau: \quad y = 0$

Latente Variable und zufällige Einflüsse

- Herleitung des Logit-Modells
- Beobachtete Variable = CDU-Wahl = $y = 1$ oder 0
- Latente Variable = „CDU-Tendenz“ = y^* = Logit
- Lineares Modell für latente Variable: $y^* = \beta_0 + \beta_1$
- Schwellenwert $\tau = 0$
  - Wenn $y^* > \tau: \quad y = 1$
  - Wenn $y^* \leqslant \tau: \quad y = 0$
- Problem: Ab bestimmtem Wert für Merkelsympathie müßten alle CDU wählen, unterhalb davon keiner

Latente Variable und zufällige Einflüsse

- Herleitung des Logit-Modells
- Beobachtete Variable = CDU-Wahl = $y = 1$ oder 0
- Latente Variable = „CDU-Tendenz“ = y^* = Logit
- Lineares Modell für latente Variable: $y^* = \beta_0 + \beta_1 + \epsilon$
- Schwellenwert $\tau = 0$
  - Wenn $y^* > \tau: \quad y = 1$
  - Wenn $y^* \leqslant \tau: \quad y = 0$
- Problem: Ab bestimmtem Wert für Merkelsympathie müßten alle CDU wählen, unterhalb davon keiner
- Zufällige Verteilung der y^*-Werte um den erwarteten Wert von y^*

Latente Variable und zufällige Einflüsse

- Herleitung des Logit-Modells
- Beobachtete Variable = CDU-Wahl = $y = 1$ oder 0
- Latente Variable = „CDU-Tendenz“ = y^* = Logit
- Lineares Modell für latente Variable: $y^* = \beta_0 + \beta_1 + \epsilon$
- Schwellenwert $\tau = 0$
  - Wenn $y^* > \tau: \quad y = 1$
  - Wenn $y^* \leqslant \tau: \quad y = 0$
- Problem: Ab bestimmtem Wert für Merkelsympathie müßten alle CDU wählen, unterhalb davon keiner
- Zufällige Verteilung der y^*-Werte um den erwarteten Wert von y^* $\rightarrow$ (konditionale) Varianz von y^*

Zufällige Einflüsse, y^* und y

- y^* hängt ab von $\beta_0 + \beta_1 x_1$ (systematische Einflüsse) und ϵ (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)

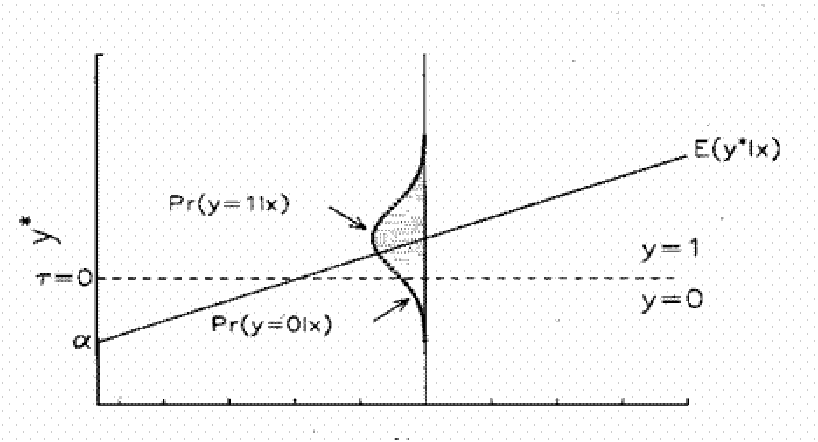
Zufällige Einflüsse, y^* und y

- 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 $y^* + \epsilon > 0 : \quad y = 1$
- Wenn $y^* + \epsilon \leqslant 0 : \quad y = 0$

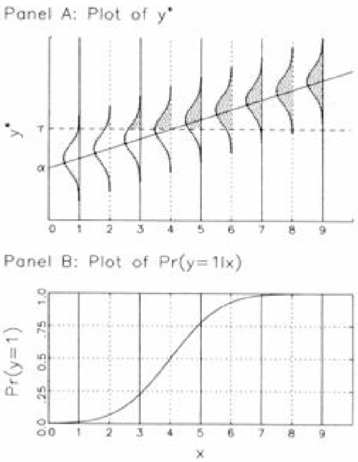
Zufällige Einflüsse, y^* und y

- 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 $y^* + \epsilon > 0 : \quad y = 1$
- Wenn $y^* + \epsilon \leqslant 0 : \quad y = 0$
- Wahrscheinlichkeit von $y = 1 =$ CDU-Wahl?

Zufällige Einflüsse, y^* und y



Zufällige Einflüsse, y^* und y

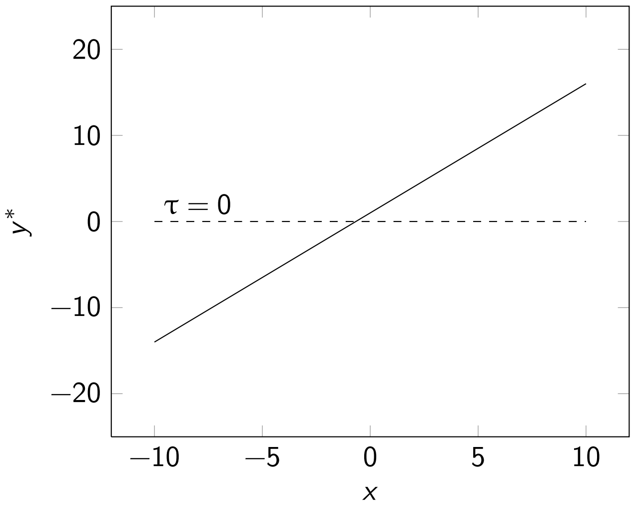


- Wahrscheinlichkeit von $y = 1$ entspricht schraffierter Fläche
- Form der Glockenkurve (konditionale Varianz von y^*) $\rightarrow$ nicht-linearer Zusammenhang zwischen Sympathie und CDU-Wahl

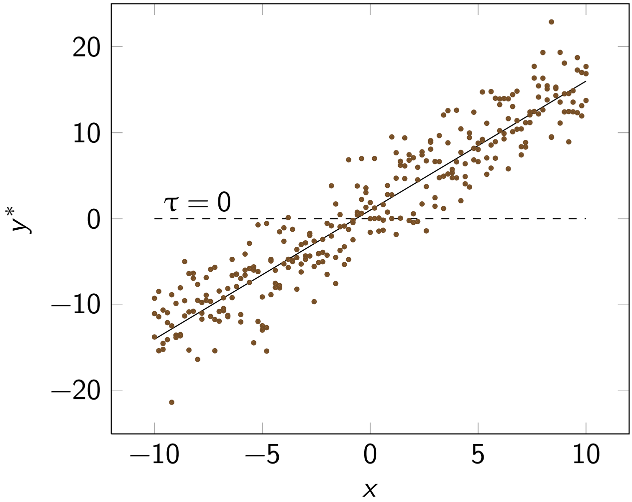
Wie war das mit der Verteilung und dem Schwellenwert?

- Latente Variable y^* – „Tendenz“ zur CDU-Wahl
- Wenn y^* Schwellenwert (0) überschreitet, wird CDU-Wahl beobachtet, ansonsten nicht
- y^* nicht vollständig durch systematischen Teil (z. B. LRS) determiniert (zufällige Einflüsse aus standardlogistischer Verteilung)
- z. B. $y^* = 1 + 1.5 \cdot x + \epsilon \rightarrow y^*$ entweder größer oder kleiner als $1.5 \cdot x$
- z. B. für $x = 0 \rightarrow 1 + 1.5 \cdot x = 1$
- y^* über oder unter Schwellenwert τ (0)?

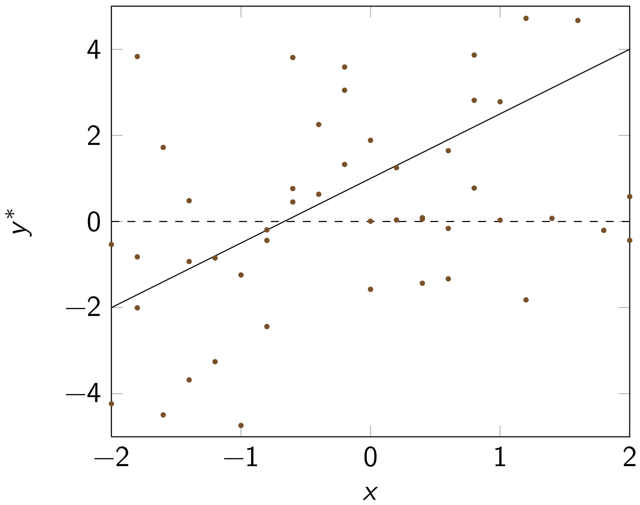
Wie war das mit der Verteilung und dem Schwellenwert?



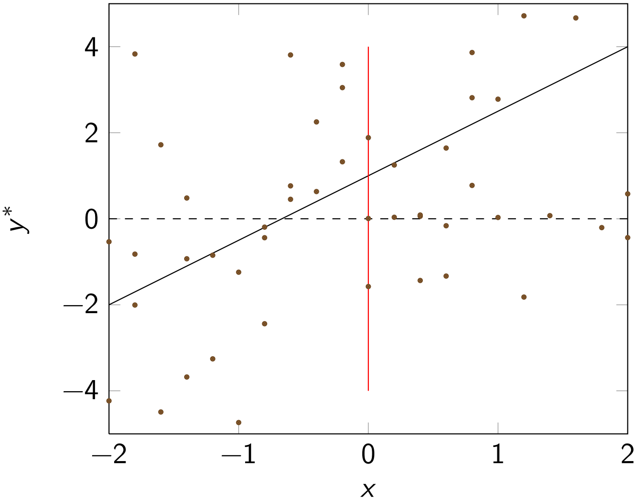
Wie war das mit der Verteilung und dem Schwellenwert?



Wie war das mit der Verteilung und dem Schwellenwert?

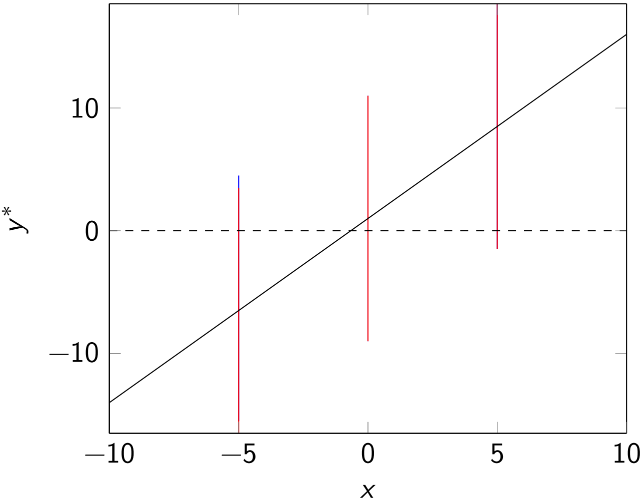


Wie war das mit der Verteilung und dem Schwellenwert?

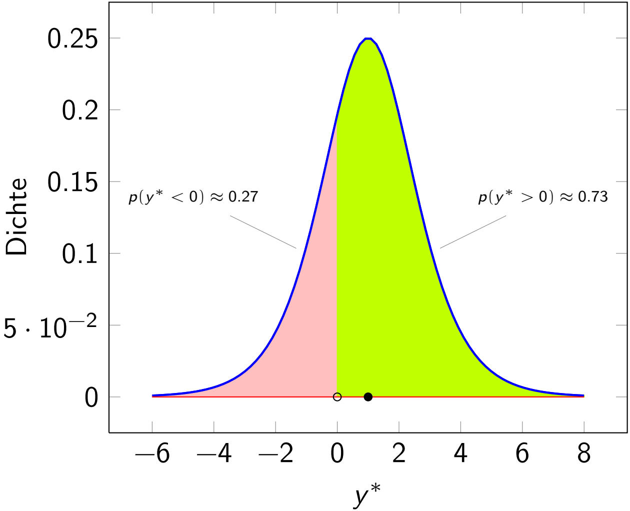


Wie wahrscheinlich

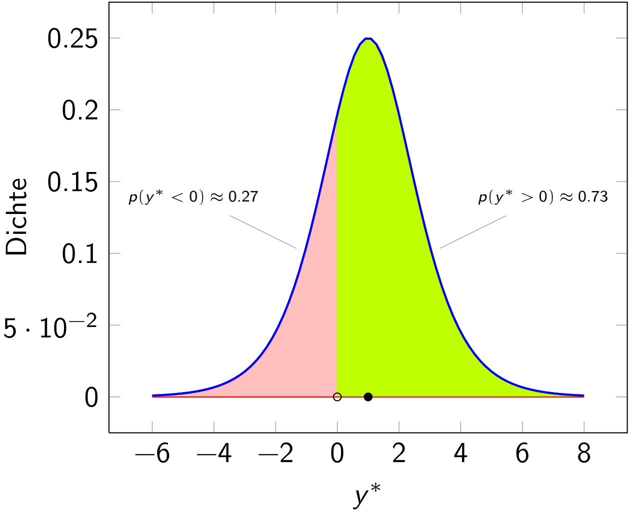
Standardlogistische Verteilung von y^*



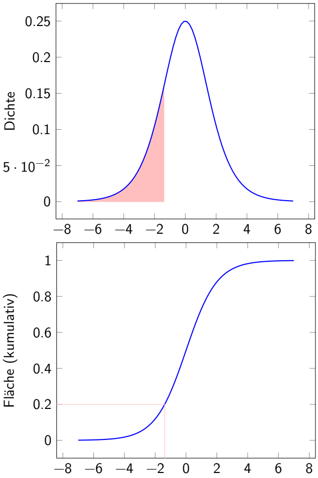
Für $x = 0 \rightarrow \mathrm{E}(y^*) = 1$



Für $x = 0 \rightarrow E(y^*) = 1$



Standardlogistische Verteilung



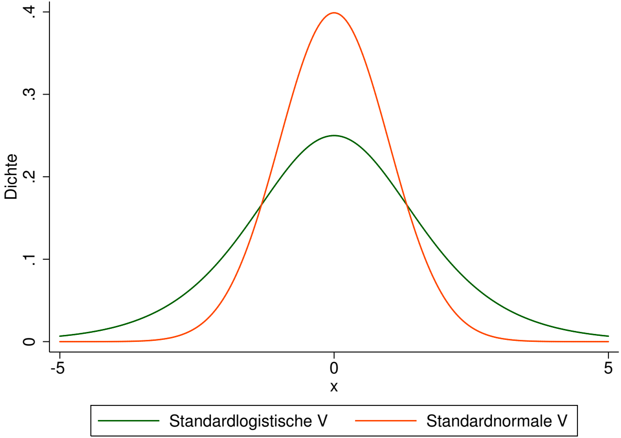
Wahrscheinlichkeitsdichte:
$f(x) = \frac{\exp(x)}{(1+\exp x)^2}$

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

Woher kennen wir Verteilung von ϵ bzw. y^*?

- y^* latent, Varianz (als Folge von ϵ) nicht beobachtbar
- *Modell-Annahme* über Form und Varianz ϵ, um Modell schätzen zu können
- Beliebte Annahme: Standardlogistische Verteilung von ϵ
  - Dichtefunktion der ϵ: $\frac{\exp(x)}{(1+\exp(x))^2}$
  - Varianz der $\epsilon = \pi^2/3 \approx 3.28$
  - Ergibt Logit-Modell
- Alternative Annahme: Standardnormalverteilung der $\epsilon \rightarrow$ Probit-Modell
- Inhaltlich normalerweise kein Unterschied, unterschiedliche Forschungstraditionen

Standardlogistische vs. Standardnormalverteilung



Ordinalskalierte Variablen

- Mehrere Ausprägungen
- Größer-Kleiner-Relation
- Keine konstanten Abstände
- Läßt sich als Schwellenwert Modell formulieren

Wieso Schwellenwerte?

- Für binäres Modell ein Schwellenwert
  $\tau = 0; y = 1 \quad \text{wenn} \quad y^* > \tau$
- 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

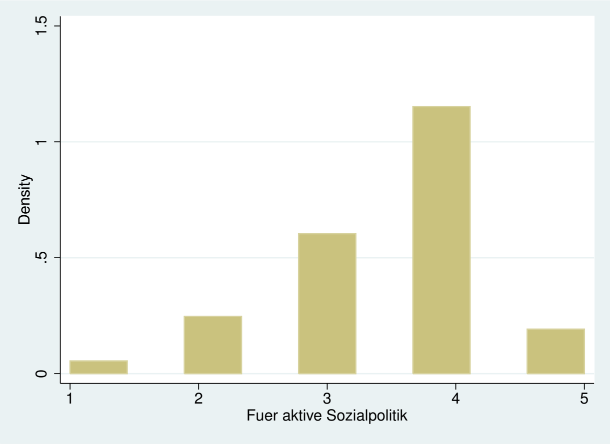
Wieso Schwellenwerte?

- Für binäres Modell ein Schwellenwert
  $\tau = 0; y = 1 \quad \text{wenn} \quad y^* > \tau$
- 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

Beispiel: Sozialpolitik

- „Soziale Sicherung sollte das wichtigste Ziel der Regierungspolitik sein“
- „Stimme überhaupt nicht zu“, „Stimme eher nicht zu“, „Weder/noch“, „Stimme eher zu“, „Stimme voll und ganz zu“

Beispiel: Sozialpolitik



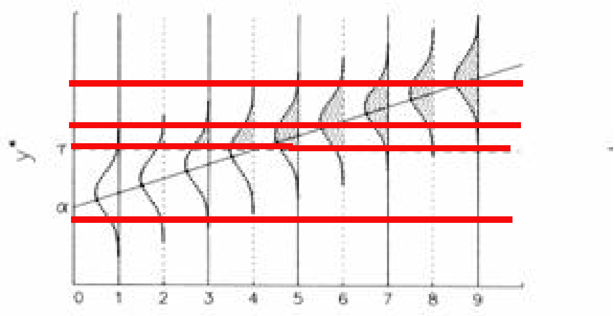
Beispiel: Sozialpolitik

- „Soziale Sicherung sollte das wichtigste Ziel der Regierungspolitik sein“
- „Stimme überhaupt nicht zu“, „Stimme eher nicht zu“, „Weder/noch“, „Stimme eher zu“, „Stimme voll und ganz zu“
- Hängt vermutlich von allgemeiner Ideologie (LRS) ab
- Modell schätzt zweierlei
  - Allgemeinen Effekt der Links-Rechts-Selbsteinstufung auf latente Variable (Logit)
  - Vier Schwellenwerte

Beispiel: Sozialpolitik

- „Soziale Sicherung sollte das wichtigste Ziel der Regierungspolitik sein"
- „Stimme überhaupt nicht zu", „Stimme eher nicht zu", „Weder/noch", „Stimme eher zu", „Stimme voll und ganz zu"
- Hängt vermutlich von allgemeiner Ideologie (LRS) ab
- Modell schätzt zweierlei
  - Allgemeinen Effekt der Links-Rechts-Selbsteinstufung auf latente Variable (Logit)
  - Vier Schwellenwerte
- Modell hat zufällige Komponente
  - Erwarteter Wert des Logit legt nicht fest, welche Ausprägung von y erwartet wird
  - Sondern wie wahrscheinlich eine Ausprägung ist

Ordered Logit mit fünf Ausprägungen



In Stata. . .

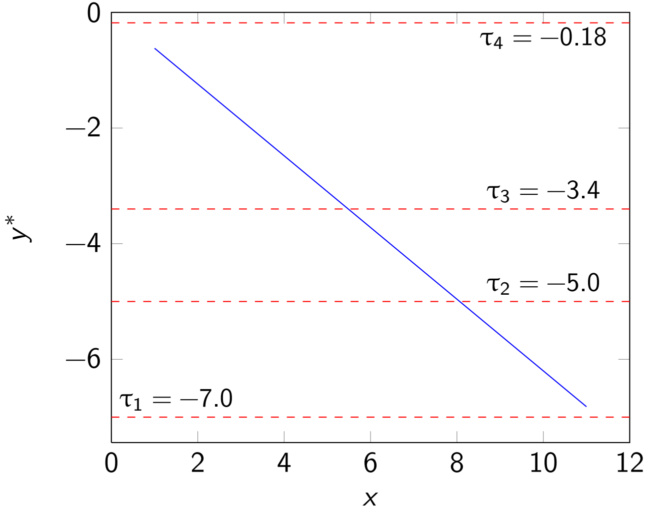
```
. ologit prosozial lrsselbstselbst

Iteration 0:   log likelihood = -98.232594
Iteration 1:   log likelihood = -88.477894
Iteration 2:   log likelihood = -88.157664
Iteration 3:   log likelihood = -88.157398
Iteration 4:   log likelihood = -88.157398

Ordered logistic regression                       Number of obs   =         79
                                                  LR chi2(1)      =      20.15
                                                  Prob > chi2     =     0.0000
Log likelihood = -88.157398                       Pseudo R2       =     0.1026

------------------------------------------------------------------------------
   prosozial |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lrsselbsts~t |  -.6180189   .1481607    -4.17   0.000    -.9084085   -.3276294
-------------+----------------------------------------------------------------
       /cut1 |  -7.041126   1.125306                     -9.246685   -4.835567
       /cut2 |  -5.040634   .8777713                     -6.761034   -3.320234
       /cut3 |   -3.42014   .7753907                     -4.939877   -1.900402
       /cut4 |  -.1760517   .6837541                     -1.516185    1.164082
------------------------------------------------------------------------------
```

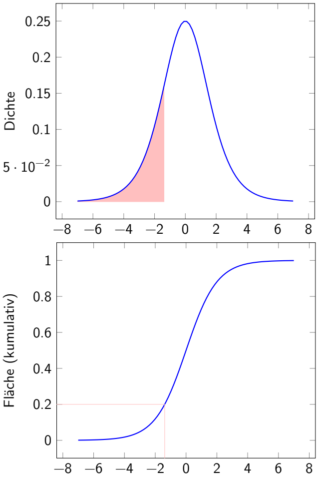
Fünf Kategorien, vier Schwellenwerte



Wahrscheinlichkeiten?

- Berechnung der Wahrscheinlichkeiten etwas komplexer als im binären Fall
- Die Wahrscheinlichkeit der Antwort „1“ ist gleich $F(\tau_1 - \mathbf{x}\boldsymbol{\beta}) = \frac{\exp(\tau_1 - \mathbf{x}\boldsymbol{\beta})}{1+\exp(\tau_1 - \mathbf{x}\boldsymbol{\beta})}$
- Die Wahrscheinlichkeit der Antwort „2“ ist gleich $F(\tau_2 - \mathbf{x}\boldsymbol{\beta}) - F(\tau_1 - \mathbf{x}\boldsymbol{\beta})$
- Die Wahrscheinlichkeit, daß $y \leqslant 2$ ist gleich $F(\tau_2 - \mathbf{x}\boldsymbol{\beta})$
- Eigentlich vier binäre Regressionen (mit identischem Parameter für LRS): 1 vs. 2,3,4,5; 1,2 vs. 3,4,5; 1,2,3 vs. 4,5; 1,2,3,4 vs. 5
- Darstellung am besten graphisch

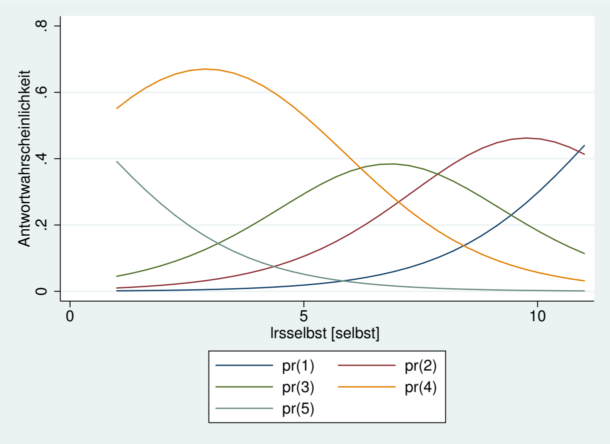
Standardlogistische Verteilung



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

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

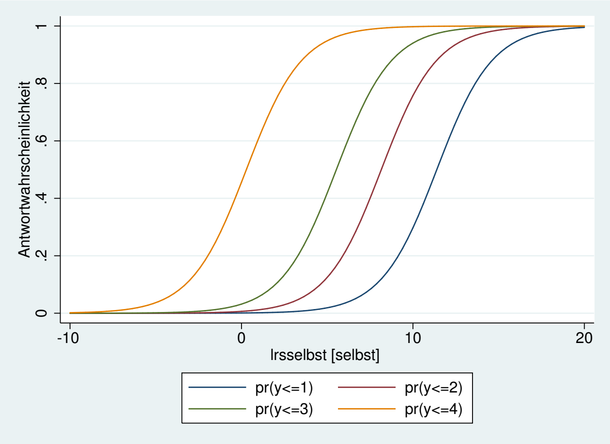
Antwortwahrscheinlichkeiten



„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

„Parallele Regression“



„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

Ordered Logit: Zusammenfassung

- Geeignet für wenige, geordnete Antwortkategorien
- Koeffizient(en) = Richtung und Signifikanz für Einfluß von $x_1 \cdots$ auf latente Variable y^*
- Zufällige (nicht-beobachtbare) Streuung der y^* um erwarteten Wert
- Schwellenwerte: Aufteilung von y^* auf beobachtbare y-Werte
- Wahrscheinlichkeit eines Wertes (z. B. „Stimme eher zu") hängt ab von
  - y^* (und damit von Koeffizient und LRS)
  - Lage der Schwellenwerte
- Antwortwahrscheinlichkeiten nicht notwendigerweise monotone Funktion von LRS (Wahrscheinlichkeit von „Stimme eher zu" nimmt erst leicht zu, fällt dann stark ab)

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

Wie sieht das multinomiale Modell aus?

- Z. B. Variable mit drei Ausprägungen (SPD (A), Union (B), Andere Entscheidung (C))
- Drei mögliche Entscheidungspaare: A vs. B, A vs. C, B vs. C
- (Grundsätzlich bei k Optionen $\frac{k \cdot (k-1)}{2}$ Entscheidungspaare)
- (Bei fünf Handlungsalternativen bereits 10 Entscheidungspaare)
- Bei *simultaner Betrachtung* ist von drei Entscheidungspaaren eines redundant
- Da 1- (Wahrscheinlichkeit(SPD) + Wahrscheinlichkeit(Union)) = Wahrscheinlichkeit(Andere)
- Eine Kategorie wird ausgelassen (Referenzkategorie)
- Trotzdem jede Menge Koeffizienten

Was drücken die Koeffizienten aus?

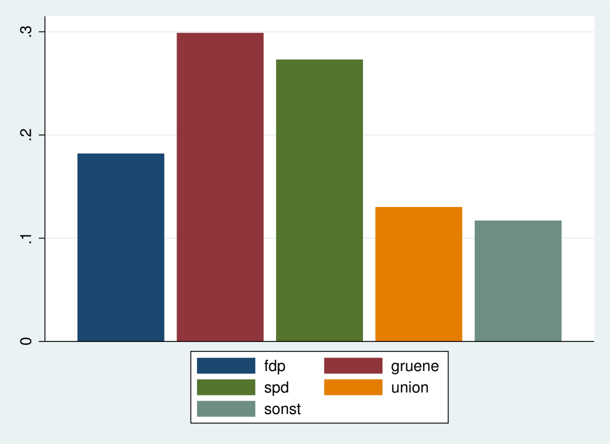
- Logit von Alternative j vs. Referenzkategorie b(Odds)
- Bzw. $\ln(\Pr(j)/\Pr(b))$
- Daraus folgt, das $\beta_b = 0$
  - Da $\frac{\Pr(b)}{\Pr(b)} = 1$ und
  - $\ln(1) = 0$
- Koeffizient für Referenzkategorie muß nicht geschätzt werden, sondern ergibt sich aus Spezifikation des Modells

Wahrscheinlichkeiten

$$\Pr(y_i = m|\mathbf{x}_i) = \frac{\exp(\mathbf{x}_i \boldsymbol{\beta}_m)}{\sum_{j=1}^{J} \exp(\mathbf{x}_i \boldsymbol{\beta}_j)}$$

- Insgesamt J (z. B. fünf) Ausprägungen
- Wahrscheinlichkeit für Ausprägung m (z. B. 4=Union) hängt ab von
  - Geschätzten Odds für diese Ausprägung (für bestimmtes Niveau von x)
  - Summe der geschätzten Odds für alle Ausprägungen (für dieses Niveau von x)
- Interpretation der Koeffizienten schwierig, unbedingt graphisch interpretieren

Beispiel Wahlabsicht und LRS



Beispiel Wahlabsicht und LRS

- Annahme: Wahlabsicht wird von Ideologie (LRS) beeinflußt
- Ganz Linke sollten vor allem „Andere“ und „Grüne“ wählen
- Ganz Rechte vermutlich FDP
- Was passiert mit CDU und SPD?
- Keine ordinale Variable, Antworten können nicht ohne weiteres zusammengefaßt werden

In Stata...

```
. mlogit wab lrsselbstselbst

Iteration 0:   log likelihood = -117.35705
Iteration 1:   log likelihood = -94.285233
Iteration 2:   log likelihood = -91.029925
Iteration 3:   log likelihood = -90.708263
Iteration 4:   log likelihood = -90.705964
Iteration 5:   log likelihood = -90.705963

Multinomial logistic regression                   Number of obs   =         76
                                                  LR chi2(4)      =      53.30
                                                  Prob > chi2     =     0.0000
Log likelihood = -90.705963                       Pseudo R2       =     0.2271

------------------------------------------------------------------------------
        wabs |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
fdp          |
lrsselbsts~t |   1.775462   .4354545     4.08   0.000     .9219871    2.628937
       _cons |  -9.651896   2.449572    -3.94   0.000    -14.45297   -4.850822
-------------+----------------------------------------------------------------
gruene       |  (base outcome)
-------------+----------------------------------------------------------------
spd          |
lrsselbsts~t |   .5628718   .2524519     2.23   0.026     .0680752    1.057668
       _cons |  -2.473407   1.101028    -2.25   0.025    -4.631382   -.3154329
-------------+----------------------------------------------------------------
union        |
lrsselbsts~t |   2.187948   .5459046     4.01   0.000     1.117995    3.257901
       _cons |  -12.71986   3.318305    -3.83   0.000    -19.22362   -6.216101
-------------+----------------------------------------------------------------
sonstige     |
lrsselbsts~t |   .0085075   .3011929     0.03   0.977    -.5818198    .5988348
       _cons |  -.9694021    1.17112    -0.83   0.408    -3.264755    1.325951
------------------------------------------------------------------------------
```

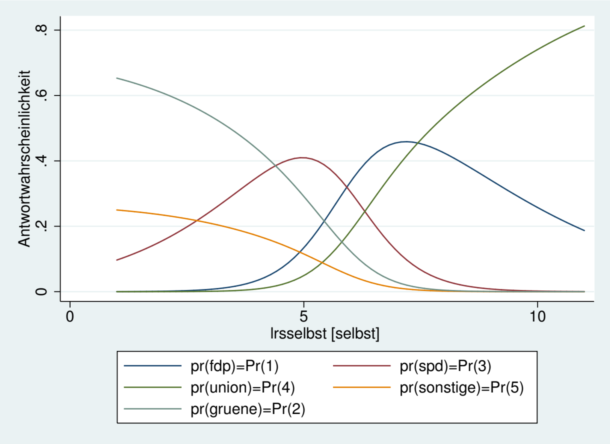
In Stata...

```
. mlogit wab lrsselbstselbst
Iteration 0:   log likelihood = -117.35705
Iteration 1:   log likelihood = -94.285233
Iteration 2:   log likelihood = -91.029925
Iteration 3:   log likelihood = -90.708263
Iteration 4:   log likelihood = -90.705964
Iteration 5:   log likelihood = -90.705963
Multinomial logistic regression                   Number of obs   =         76
                                                  LR chi2(4)      =      53.30
                                                  Prob > chi2     =     0.0000
Log likelihood = -90.705963                       Pseudo R2       =     0.2271
```

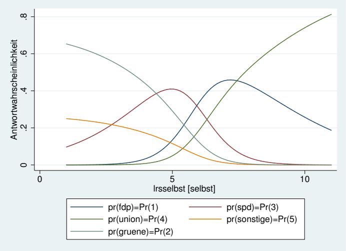
wabs	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
fdp						
lrsselbsts~t	1.775462	.4354545	4.08	0.000	.9219871	2.628937
_cons	-9.651896	2.449572	-3.94	0.000	-14.45297	-4.850822
gruene	(base outcome)					
spd						
lrsselbsts~t	.5628718	.2524519	2.23	0.026	.0680752	1.057668
_cons	-2.473407	1.101028	-2.25	0.025	-4.631382	-.3154329
union						
lrsselbsts~t	2.187948	.5459046	4.01	0.000	1.117995	3.257901
_cons	-12.71986	3.318305	-3.83	0.000	-19.22362	-6.216101
sonstige						
lrsselbsts~t	.0085075	.3011929	0.03	0.977	-.5818198	.5988348
_cons	-.9694021	1.17112	-0.83	0.408	-3.264755	1.325951

- Recht hohes Pseudo-R^2
- Grüne = Basiskategorie (beliebig, ändert nichts am Modellfit)
- LRS unterscheidet nicht signifikant zwischen Grünen- und Anderen-Wählern
- Signifikante Effekte für Effekt LRS auf Wahl von SPD, FDP, Union vs. Grüne
- Effekte positiv – höhere Wahrscheinlichkeiten für rechtere Wähler
- Erweiterung zum multivariaten Modell möglich (z. B. mit Merkelsympathie etc.)

LRS und erwartetes Wahlverhalten



LRS und erwartetes Wahlverhalten



- Komplexe, nicht-lineare Zusammenhänge
- Summe der erwarteten Wahrscheinlichkeiten an jedem Punkt der LRS =1
- Grüne/andere im Mitte-Links Bereich populär, rechts der Mitte rasanter Abfall

- (Sehr gutes Grünen-Ergebnis weil Gruppe relativ links)
- SPD im ideologischen Zentrum und links der Mitte erfolgreich (wenn auch nicht sehr)
- Im Bereich von 5–7 FDP erfolgreicher als Union, dann rasanter Abfall/Anstieg

Was gibt es sonst noch?

- Flexiblere Modelle für ordinale Daten
- Alternativenspezifische Modelle für kategoriale Daten (Eigenschaften des Wählers und der Partei)
- ...

Was gibt es sonst noch?

- Flexiblere Modelle für ordinale Daten
- Alternativenspezifische Modelle für kategoriale Daten (Eigenschaften des Wählers und der Partei)
- ...
- ... nicht in dieser Vorlesung

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