

Übungsaufgabe 6

Fabio Votta

12. Dezember 2017

Aufgabe 1

Erstellen Sie zunächst dichotome Variablen für SPD-Wahl bei der letzten Bundestagswahl (z. B. V810), die Gewerkschaftsmitgliedschaft (z. B. V606) und die ev. Konfessionszugehörigkeit (z. B. V601).

V810 ZWEITSTIMME, LETZTE BUNDESTAGSWAHL

- 1 CDU-CSU
- 2 SPD
- 3 FDP
- 4 DIE GRUENEN
- 6 DIE LINKE
- 20 NPD
- 41 PIRATEN
- 42 AFD
- 43 FREIE WAEHLER
- 90 ANDERE PARTEI

V606 MITGLIED IN EINER GEWERKSCHAFT?

- 1 Ja
- 2 Nein

V601 KONFESSION, BEFRAGTE

- 1 EVANG.OHNE FREIKIRCH
- 2 EVANG.FREIKIRCHE
- 3 ROEMISCH-KATHOLISCH
- 4 AND.CHRISTL.RELIGION
- 5 AND.NICHT-CHRISTLICH
- 6 KEINER RELIGIONSGEM.

Tipp: Benutze zunächst `select` um alle benötigten Variablen auszuwählen, dann kodiere die Dummies.

```
allbus %<>%
  select(V810, V606, V601) %>%
  mutate(spd = ifelse(V810 == 2,1,0)) %>%
  mutate(gewerkschaft = ifelse(V606 == 1, 1, 0)) %>%
  mutate(evangelisch = ifelse(V601 == 1 | V601 == 2, 1,0)) %>%
  mutate(katholisch = ifelse(V601 == 3, 1, 0))
```

Aufgabe 1a_i

Berechnen Sie (anhand eines log. Regressionsmodells) die Wahrscheinlichkeit für ein aktuelles Gewerkschaftsmitglied (Konfession evangelisch) die SPD zu wählen.

```
mod1 <- glm(spd ~ gewerkschaft + evangelisch, data = allbus, family = "binomial")  
  
texreg(mod1, float.pos = "!h")
```

	Model 1
(Intercept)	-1.27*** (0.09)
gewerkschaft	0.37 (0.20)
evangelisch	0.50*** (0.14)
AIC	1322.06
BIC	1337.22
Log Likelihood	-658.03
Deviance	1316.06
Num. obs.	1157

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 1: Statistical models

```
mod1_wide <- tidy_wide(mod1)  
  
mod1_wide %>%  
  mutate(z = intercept + 1 * gewerkschaft + 1 * evangelisch) %>%  
  mutate(p = logit2prob(z)) %>%  
  kable()
```

intercept	evangelisch	gewerkschaft	z	p
-1.267753	0.4955968	0.3653885	-0.406768	0.3996873

Die Wahrscheinlichkeit für ein evangelisches, aktuelles Gewerkschaftsmitglied, die SPD zu wählen liegt bei knapp 40 Prozent.

Aufgabe 1a_{ii}

Berechnen Sie (anhand eines log. Regressionsmodells) die Wahrscheinlichkeit für eine Person mit katholischer Konfession (keine Gewerkschaftsmitgliedschaft) die SPD zu wählen.

```
mod2 <- glm(spd ~ gewerkschaft + katholisch, data = allbus, family = "binomial")  
  
texreg(mod2, float.pos = "!h")
```

	Model 1
(Intercept)	-1.00*** (0.08)
gewerkschaft	0.36 (0.19)
katholisch	-0.32* (0.16)
AIC	1330.93
BIC	1346.09
Log Likelihood	-662.46
Deviance	1324.93
Num. obs.	1157

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 3: Statistical models

```
mod2_wide <- tidy_wide(mod2)  
  
mod2_wide %>%  
  mutate(z = intercept + 0 * gewerkschaft + 1 * katholisch) %>%  
  mutate(p = logit2prob(z)) %>%  
  kable()
```

intercept	gewerkschaft	katholisch	z	p
-1.002159	0.3600558	-0.3218587	-1.324018	0.2101506

Die Wahrscheinlichkeit für ein katholisches Nicht-Gewerkschaftsmitglied die SPD zu wählen liegt bei 21.02 Prozent.

Aufgabe 1b

Wie fallen die Odds für diese beiden Personen aus?

```
mod1_wide %>%  
  mutate(z = intercept + 1 * gewerkschaft + 1 * evangelisch) %>%  
  mutate(odds = 1 / exp(z)) %>%  
  kable()
```

intercept	evangelisch	gewerkschaft	z	odds
-1.267753	0.4955968	0.3653885	-0.406768	1.501956

```
mod2_wide %>%
  mutate(z = intercept + 0 * gewerkschaft + 1 * katholisch) %>%
  mutate(odds = 1 / exp(z)) %>%
  kable()
```

intercept	gewerkschaft	katholisch	z	odds
-1.002159	0.3600558	-0.3218587	-1.324018	3.758493

Die Odds für eine Person (A), die sowohl evangelisch, als auch aktuell Gewerkschaftsmitglied ist, beträgt: 0.67 Das heißt: Damit ist die Chance, dass diese Person (A) SPD statt einer anderen Partei wählt 0.67 mal höher oder 1.50 mal niedriger.

Die Odds für eine Person (B), die katholisch und aktuell kein Gewerkschaftsmitglied ist, beträgt: 3,553 Das heißt: Damit ist die Chance, dass diese Person (B) SPD statt einer anderen Partei wählt 3,76 mal niedriger.

Aufgabe 1c

Wie fallen die Odds Ratios bzgl. der Variablen Gewerkschaftsmitgliedschaft und Konfessionszugehörigkeit-evangelisch aus?

```
tidy(mod1) %>%
  mutate(odds_ratio = exp(estimate)) %>%
  kable()
```

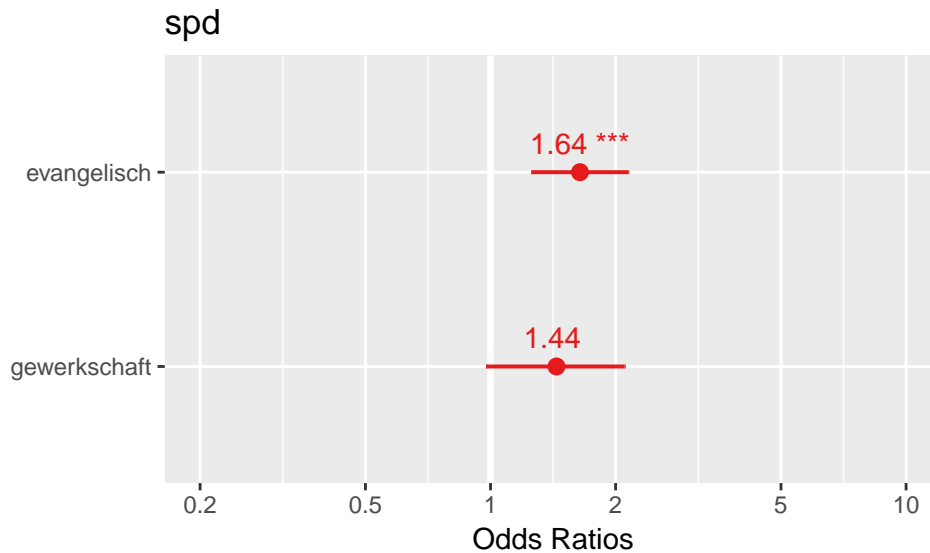
term	estimate	std.error	statistic	p.value	odds_ratio
(Intercept)	-1.2677532	0.0912714	-13.889924	0.0000000	0.2814633
gewerkschaft	0.3653885	0.1953941	1.870007	0.0614828	1.4410737
evangelisch	0.4955968	0.1369813	3.617988	0.0002969	1.6414775

Im Vergleich zu Personen, die nicht aktuell Gewerkschaftsmitglied sind, haben Gewerkschaftsmitglieder eine 1,44 mal höhere Chance die SPD zu wählen.

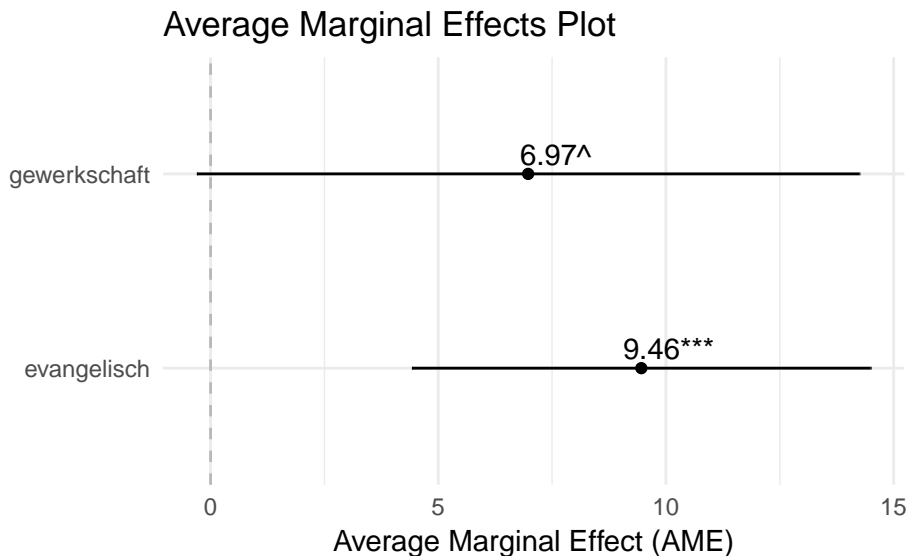
Im Vergleich zu Personen, die keine evangelische Konfessionsangehörigkeit haben, haben Personen mit evangelischer Zugehörigkeit eine 1,64 mal höhere Chance die SPD zu wählen.

Visualisierung für Bonuspoints ;)

```
plot_model(mod1, show.values = T, show.p = T)
```



```
plot_ame(mod1)
```



Aufgabe 1d

Über welchen Personenkreis kann dieses Modell Aussagen treffen? Denken Sie bei Ihrer Antwort an die Null-Kategorie Ihrer Variablen.

Das Modell kann nur Aussagen über Personen mit evangelischer Konfessionszugehörigkeit und Personen ohne evangelische Konfessionszugehörigkeit machen. Das heißt, es kann —keine— Aussagen über Atheisten, Muslime etc. machen. Es lässt sich auch nichts über die Wahrscheinlichkeit/Chance sagen eine spezielle Partei zu wählen (wie z.B. CDU oder Grüne). Die Dummy Variable ist nur als 1 (SPD Wahl) und 0 (alle anderen Parteien) kodiert.