

Statistik II

Einheit 11: Verfahren für Nominaldaten (2)

03.07.2024 | Prof. Dr. Stephan Goerigk

Logistische Regression

Kurzvorstellung:

- In allen bisher besprochenen Regressionsmodellen war die abhängige Variable (AV) eine stetige Variable.
- Ziel: Regressionsmodelle mit diskreter AV
- Vorteile gegenüber χ^2 -Tests
 - Vorhersage mittels Prädiktoren
 - Erweiterbarkeit auf **multiple logistische Regression**
 - Testen von Interaktionen zwischen 2 Prädiktoren
- Logistische Regression bringt Logik und Vorteile anderer Regressionen mit sich

Logistische Regression

Beispiel:

- AV : Erfolg einer Gesprächstherapie (dummy-codiert)
 - $Y_i = 0$ falls nein
 - $Y_i = 1$ falls ja
- UV_1 : (stetige) gemessene Empathie der Therapeut:in
- UV_2 : (diskreter) erhobener Bindungsstil der Patient:in

Frage: Wie lässt sich der Zusammenhang zwischen den Variablen beschreiben?

Beispiel - nominalskalierte (diskrete) AV:

Behandlungserfolg	Empathie	Bindungsstil
ja	8	sicher
ja	9	sicher
nein	4	unsicher
ja	15	sicher
ja	10	sicher
nein	5	unsicher
nein	10	unsicher
nein	7	unsicher
ja	10	sicher
ja	15	sicher
nein	7	unsicher
nein	6	unsicher
nein	4	unsicher
nein	11	sicher
ja	11	sicher
nein	7	unsicher
ja	6	sicher

Logistische Regression

Wir beginnen mit einem bivariaten Modell (AV und UV_1) :

- AV : Erfolg einer Gesprächstherapie (dummy-codiert)
 - $Y_i = 0$ falls nein
 - $Y_i = 1$ falls ja
 - UV_1 : (stetige) gemessene Empathie der Therapeut:in
-

Frage: Wie lässt sich der Zusammenhang zwischen den Variablen beschreiben?

Lösungsvorschlag: Einfaches lineares Regressionsmodell (bereits bekannt)

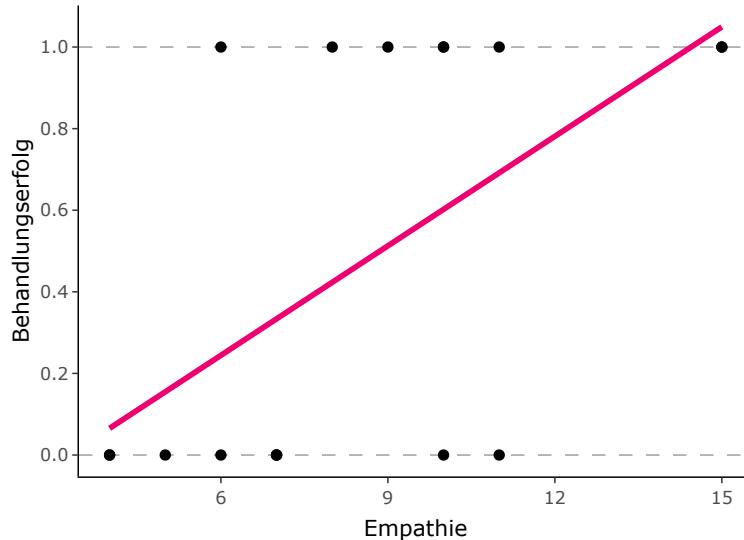
$$Y_i = a + \beta \cdot x_i + \varepsilon_i$$

Problem:

- Die meisten der mit der Gerade vorhergesagten Werte können gar nicht tatsächlich beobachtet werden
- Lediglich Werte 0 und 1 können für die dummy-codierte Variable auftreten

Logistische Regression

Lösungsvorschlag: Einfaches lineares Regressionsmodell



$$Y_i = a + \beta \cdot x_i + \varepsilon_i$$

- Modellierung mittels Y-Achsenabschnitt (a) und Steigung (β)
- Gerade zeigt positiven Zusammenhang zwischen Empathie und Therapieerfolg
 - je mehr Empathie desto besser die Therapie
 - Dies erscheint zunächst plausibel
- ABER: Modell ignoriert Begrenzung des numerischen Bereichs der AV auf Werte 0 und 1 (Misserfolg vs. Erfolg)

Logistische Regression

Modell für bedingte Wahrscheinlichkeiten:

- In statistischen Modellen für binäre AVs wird i.d.R. die (bedingte) **Wahrscheinlichkeit** (P) modelliert, dass Y_i die Werte 0 bzw. 1 annimmt.
 - Wahrscheinlichkeit (P) anstelle der Werte von Variable Y_i selbst
 - Wahrscheinlichkeit ist "bedingt" bei gegebenen Werten von Prädiktoren $x_{i1}, \dots x_{ik}$
- Wenn 0 und 1 komplementäre Ereignisse sind, reicht es aus nur $P(Y_i = 1|x_{i1}, \dots x_{ik})$ zu modellieren, denn

$$P(Y_i = 1|x_{i1}, \dots x_{ik}) = 1 - P(Y_i = 0|x_{i1}, \dots x_{ik})$$

Logistische Regression

Lineares Wahrscheinlichkeitsmodell

In unserem Beispiel:

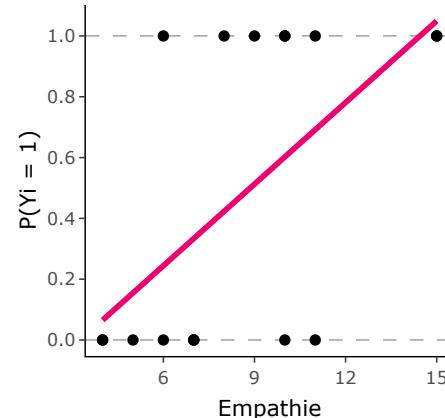
- AV : Erfolg einer Gesprächstherapie (dummy-codiert)
 - $Y_i = 0$ falls nein
 - $Y_i = 1$ falls ja
- UV_1 : (stetige) gemessene Empathie der Therapeut:in

Wahrscheinlichkeit für Behandlungserfolg $P(Y_i = 1|x_i)$ wird mit Geradengleichung vorhergesagt:

$$P(Y_i = 1|x_i) = a + \beta \cdot x_i + \varepsilon$$

ABER: Das Problem bleibt bestehen

- Wahrscheinlichkeit liegt zwischen 0 und 1
- Wertebereich von $a + \beta \cdot x_i + \varepsilon$ liegt zwischen $-\infty$ und $+\infty$



Logistische Regression

Die logistische Funktion

Ziel: Transformation des Wertebereichs $-\infty$ bis $+\infty$ in Werte zwischen 0 und 1 (Wertebereich von Wahrscheinlichkeiten)

Modellgleichung der einfachen logistischen Regression:

$$P(Y_i = 1|x_i) = \frac{e^{a+\beta \cdot x_i}}{1 + e^{a+\beta \cdot x_i}}$$

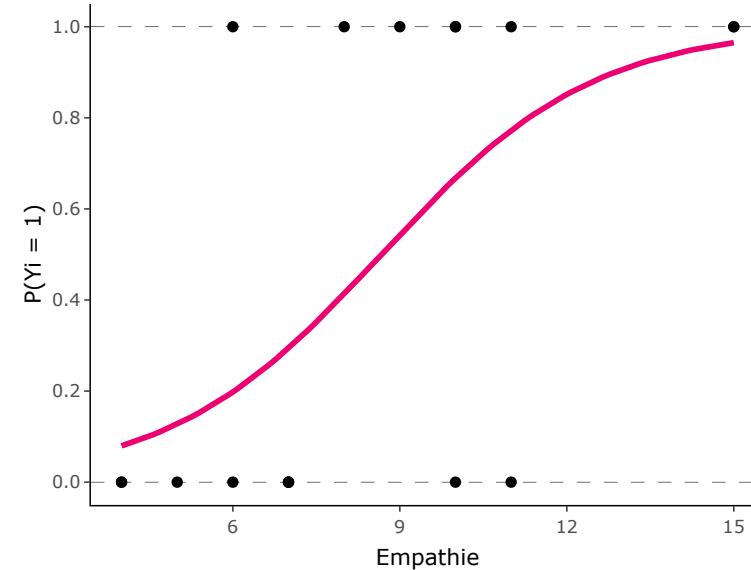
- Da wir nicht Y_i selbst, sondern die bedingte Wahrscheinlichkeit für $Y_i = 1$ modellieren, benötigen wir keinen Fehlerterm

Logistische Regression

Die logistische Funktion

$$P(Y_i = 1|x_i) = \frac{e^{a+\beta \cdot x_i}}{1 + e^{a+\beta \cdot x_i}}$$

- Mit Anwendung der logistischen Funktion auf die lineare Regression lässt sich die Gerade "biegen"
- Sie ist nun auf den Wertebereiche zwischen 0 und 1 begrenzt
- Am Wendepunkt der Kurve ist $P(Y_i = 1|x_i) = 0.5$
- Beide Optionen (Erfolg vs. Misserfolg) sind hier gleich wahrscheinlich



Logistische Regression

Modellparameter

$$P(Y_i = 1|x_i) = \frac{e^{a+\beta \cdot x_i}}{1 + e^{a+\beta \cdot x_i}}$$

- Parameter β (entspricht dem Steigungsparameter):
 - $\beta = 0$: Wahrscheinlichkeit für $Y_i = 1$ ist unabhängig von der UV (kein Zusammenhang).
 - $\beta > 0$: Je höher der Wert auf der UV, desto höher die Wahrscheinlichkeit für $Y_i = 1$.
 - $\beta < 0$: Je höher der Wert auf der UV, desto niedriger die Wahrscheinlichkeit für $Y_i = 1$.
- Parameter a (entspricht dem Y-Achsenabschnitt)

Logistische Regression

Interpretation der Modellparameter - Odds

- Der Bruch aus der Wahrscheinlichkeit für ein Ereignis geteilt durch dessen Gegenwahrscheinlichkeit, wird als **Odds** bezeichnet
- In der logistischen Regression relevante Odds:

$$\frac{P(Y_i = 1|x_i)}{P(Y_i = 0|x_i)}$$

- Falls das Ereignis $Y_i = 1$ wahrscheinlicher ist als das Ereignis $Y_i = 0 \rightarrow \text{Odds} > 1$
- Falls das Ereignis $Y_i = 1$ wahrscheinlicher ist als das Ereignis $Y_i = 0 \rightarrow \text{Odds} < 1$
- Sind beide Ereignisse gleich wahrscheinlich $\rightarrow \text{Odds} = 1$

Logistische Regression

Ermittlung der Odds für Steigung (β)

Odds lassen sich durch Einsetzen der Modellparameter (a, β) in die Odds Formel ermitteln:

$$\frac{P(Y_i = 1|x_i)}{P(Y_i = 0|x_i)} = \frac{P(Y_i = 1|x_i)}{1 - P(Y_i = 1|x_i)} = e^a \cdot e^{\beta \cdot x_i}$$

Das heißt: Erhöht sich die UV um eine Einheit, erhöhen sich die Odds um den Faktor e^β

- $e^\beta = 1$: kein Einfluss der UV auf die Odds
- $e^\beta > 1$: wenn Wert der UV höher wird, werden auch Odds höher
- $e^\beta < 1$: wenn Wert der UV höher wird, werden Odds niedriger

Logistische Regression

Ermittlung der Odds für Y-Achsenabschnitt (β)

Wenn $x_i = 0$ gilt für die Odds

$$\frac{P(Y_i = 1|x_i = 0)}{P(Y_i = 0|x_i = 0)} = e^a \cdot e^{\beta \cdot 0} = e^a \cdot e^0 = e^a$$

- Für Personen mit einem Wert von 0 auf der UV, sind die Odds gleich e^a
- Kennen wir bereits von linearer Regression: Y-Achsenabschnitt ist Wert, wenn $x_i = 0$ (UV = 0)
- Auch hier hängt Sinnhaftigkeit der Interpretation des Werts davon ab, ob ein UV-Wert von 0 sinnvoll interpretierbar ist (s.h. z.B. IQ)
- Falls UV-Wert von 0 nicht sinnvoll interpretierbar → **Zentrierung der UV** (s.h. lineare Regression)

Logistische Regression

Einheit von a und β (Log-Odds)

- Mittels der Odds haben wir also eine inhaltliche Interpretation der Modellparameter
- Wie gesehen, müssen diese dafür jeweils zunächst mittels Euler'scher Zahl e exponiert werden (e^a bzw. e^β)

ABER: In welcher Einheit werden a und β selbst dann gemessen und warum ist das Exponieren notwendig?

- Nimmt man von den Odds den natürlichen Logarithmus, so erhält man die „**Log-Odds**“ bzw. „**Logits**“:

$$\ln\left(\frac{P(Y_i = 1|x_i)}{P(Y_i = 0|x_i)}\right)$$

- Wenn $P(Y_i = 1|x_i) = 0.5 \rightarrow \text{Log-Odds} = 0$
- Wenn $P(Y_i = 1|x_i) < P(Y_i = 0|x_i) \rightarrow \text{Log-Odds} < 0$
- Wenn $P(Y_i = 1|x_i) > P(Y_i = 0|x_i) \rightarrow \text{Log-Odds} > 0$

Logistische Regression

Einheit von a und β (Log-Odds)

- Einsetzen der Modellgleichung für die Odds in die Definition der Log-Odds ergibt:

$$\ln\left(\frac{P(Y_i = 1|x_i)}{P(Y_i = 0|x_i)}\right) = \ln(e^{a+\beta \cdot x_i}) = a + \beta \cdot x_i$$

→ Logistische Regression ist eine **lineare Regression der Log-Odds**

Die Parameter a und β können damit wie folgt interpretiert werden:

- a entspricht den Log-Odds für Personen mit einem UV-Wert von 0
 - (wie Y-Achsenabschnitt linearer Regression).
- Wenn sich die UV um eine Einheit erhöht, erhöhen sich die Log-Odds um β
 - (wie Steigung linearer Regression).
- Da die Log-Odds weniger intuitiv als die Odds sind, werden jedoch **meist e^a und e^β (Odds) statt a und β (Log-Odds) interpretiert.**

Logistische Regression

Zurück zum Beispiel:

- AV : Erfolg einer Gesprächstherapie (dummy-codiert)
 - $Y_i = 0$ falls nein
 - $Y_i = 1$ falls ja
- UV_1 : (stetige) gemessene Empathie der Therapeut:in
- UV_2 : (diskreter) erhobener Bindungsstil der Patient:in

Frage: Wie lässt sich der Zusammenhang zwischen den Variablen beschreiben?

Beispiel - nominalskalierte (diskrete) AV:

Behandlungserfolg	Empathie	Bindungsstil
ja	8	sicher
ja	9	sicher
nein	4	unsicher
ja	15	sicher
ja	10	sicher
nein	5	unsicher
nein	10	unsicher
nein	7	unsicher
ja	10	sicher
ja	15	sicher
nein	7	unsicher
nein	6	unsicher
nein	4	unsicher
nein	11	sicher
ja	11	sicher
nein	7	unsicher
ja	6	sicher

Logistische Regression

Zurück zum Beispiel:

$$P(Y_i = 1|x_i) = \frac{e^{a+\beta \cdot x_i}}{1 + e^{a+\beta \cdot x_i}}$$

Gedacht in Odds:

- Die Odds für einen Behandlungserfolg bei einer Therapeut:in mit Empathie von 0 sind e^a
- Falls sich die Empathie um 1 Einheit erhöht, erhöhen sich die Odds für den Behandlungserfolg um den Faktor e^β

Gedacht in Log-Odds:

- Die Log-Odds für einen Behandlungserfolg bei einer Therapeut:in mit Empathie von 0 sind a
- Falls sich die Empathie um 1 Einheit erhöht, erhöhen sich die Log-Odds für den Behandlungserfolg (linear) um β

Beispiel - nominalskalierte (diskrete) AV:

Behandlungserfolg	Empathie	Bindungsstil
ja	8	sicher
ja	9	sicher
nein	4	unsicher
ja	15	sicher
ja	10	sicher
nein	5	unsicher
nein	10	unsicher
nein	7	unsicher
ja	10	sicher
ja	15	sicher
nein	7	unsicher
nein	6	unsicher
nein	4	unsicher
nein	11	sicher
ja	11	sicher
nein	7	unsicher
ja	6	sicher

Logistische Regression

- In R benutzen wir statt `lm` nun `glm` (generalized linear model)

```
model = glm(Behandlungserfolg ~ Empathie, data = data, family = "binomial")
summary(model)

##
## Call:
## glm(formula = Behandlungserfolg ~ Empathie, family = "binomial",
##      data = data)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.5438     2.2836 -1.990   0.0466 *
## Empathie     0.5242     0.2644  1.983   0.0474 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 23.508 on 16 degrees of freedom
## Residual deviance: 16.802 on 15 degrees of freedom
## AIC: 20.802
##
## Number of Fisher Scoring iterations: 5
```

- Output zeigt unter **Estimate** immer a und β (Log-Odds)
- Diese folgen der Logik des linearen Modells

Umwandlung Log-Odds in Odds mittels Exponierung (z.B. e^β):

```
exp(-4.5438)
## [1] 0.01063292
exp(0.5242)
## [1] 1.689107
```

- $0.01 = \text{Odds für Behandlungserfolg, wenn Empathie} = 0$
- Zunahme der Odds um Faktor 1.69, wenn Empathie um 1 Einheit zunimmt

Logistische Regression

- In R benutzen wir statt `lm` nun `glm` (generalized linear model)

```
model = glm(Behandlungserfolg ~ Empathie, data = data, family = "binomial")
summary(model)

##
## Call:
## glm(formula = Behandlungserfolg ~ Empathie, family = "binomial",
##      data = data)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.5438     2.2836 -1.990   0.0466 *
## Empathie     0.5242     0.2644  1.983   0.0474 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 23.508 on 16 degrees of freedom
## Residual deviance: 16.802 on 15 degrees of freedom
## AIC: 20.802
##
## Number of Fisher Scoring iterations: 5
```

Hypothesentests:

- $\Pr(>|z|)$ zeigt an, das ein Hypothesentest gerechnet wurde

Dieser prüft folgende Hypothesenpaare:

- Für a
 - $H_0 : a = 0$
 - $H_1 : a \neq 0$
- Für β
 - $H_0 : \beta = 0$
 - $H_1 : \beta \neq 0$
- $p_\beta = 0.0474 < .05$

→ Wir gehen davon aus, dass es einen signifikanten Zusammenhang zwischen Empathie und Therapieerfolg gibt.

Logistische Regression

- In R benutzen wir statt `lm` nun `glm` (generalized linear model)

```
model = glm(Behandlungserfolg ~ Empathie, data = data, family = "binomial")
car::Anova(model)

## Analysis of Deviance Table (Type II tests)
##
## Response: Behandlungserfolg
##           LR Chisq Df Pr(>Chisq)
## Empathie    6.706  1   0.009609 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Hypothesentests (Omnibustest):

- Auch für Log-Regression lässt sich ein Omnibustest rechnen
- Dieser prüft, ob Hinzunahme des Prädiktors Modellpassung signifikant verbessert (Modell mit Prädiktor mehr Varianz erklärt)
- Prüft *nicht* einzelne Steigungsparameter
- Die Teststatistik ist χ^2 -verteilt

Logistische Regression

Berechnung der Konfidenzintervalle für Log-Odds und Odds:

```
model = glm(Behandlungserfolg ~ Empathie, data = data, family = "binomial")  
  
# Für Log-Odds  
confint(model)
```

```
##                2.5 %    97.5 %  
## (Intercept) -10.2498495 -0.8737261  
## Empathie      0.1066026  1.1812475
```

```
# Für Odds  
exp(confint(model))
```

```
##                2.5 %    97.5 %  
## (Intercept) 0.00003536282 0.4173934  
## Empathie     1.11249211601 3.2584367
```

- Wir gehen davon aus, dass Odds für einen Behandlungserfolg um Faktor 1.11 bis 3.26 steigen, wenn Empathie um 1 Einheit zunimmt
- Gängige Notation (engl.): $OR = 1.69, CI_{95\%} = 1.11 \text{ to } 3.26$

Logistische Regression mit zentrierter UV

```
model = glm(Behandlungserfolg ~ center(Empathie), data = data, family = "binomial")
summary(model)

##
## Call:
## glm(formula = Behandlungserfolg ~ center(Empathie), family = "binomial",
##      data = data)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.07312   0.60264 -0.121  0.9034
## center(Empathie)  0.52415   0.26435  1.983  0.0474 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 23.508 on 16 degrees of freedom
## Residual deviance: 16.802 on 15 degrees of freedom
## AIC: 20.802
##
## Number of Fisher Scoring iterations: 5
```

- Zentrierung (grand mean): Von jedem Wert der UV wird der Durchschnitt der UV abgezogen
- Empathie (zentriert) = 0 ist nun der Durschnittswert der Stichprobe (nur α ändert sich, β bleibt unverändert)

Umwandlung Log-Odds in Odds mittels Exponierung (z.B. e^β):

```
exp(-0.07)
## [1] 0.9323938

exp(0.5242)
## [1] 1.689107
```

- 0.93 = Odds für Behandlungserfolg **bei durchschnittlicher Empathie**
- Zunahme der Odds um Faktor 1.69, wenn Empathie um 1 Einheit zunimmt

Multiple logistische Regression

- Wie bei der linearen Regression lässt sich auch die einfache logistische Regression zur multiplen logistischen Regression erweitern:

$$P(Y_i = 1|x_{i1}, \dots, x_{ik}) = \frac{e^{a + \beta_1 \cdot x_i + \dots + \beta_k \cdot x_k}}{1 + e^{a + \beta_1 \cdot x_i + \dots + \beta_k \cdot x_k}}$$

Gedacht in Odds:

$$\frac{P(Y_i = 1|x_{i1}, \dots, x_{ik})}{P(Y_i = 0|x_{i1}, \dots, x_{ik})} = e^{a + \beta_1 \cdot x_i + \dots + \beta_k \cdot x_k}$$

Gedacht in Log-Odds:

$$\ln\left(\frac{P(Y_i = 1|x_{i1}, \dots, x_{ik})}{P(Y_i = 0|x_{i1}, \dots, x_{ik})}\right) = a + \beta_1 \cdot x_i + \dots + \beta_k \cdot x_k$$

Multiple logistische Regression

Interpretation der Modellparameter:

Gedacht in Odds:

- e^a entspricht den Odds von Personen (z.B. für Behandlungserfolg), die auf allen UVs den Wert 0 haben.
- e^{β_j} entspricht dem Faktor, um den sich die Odds erhöhen, falls sich die UV j um eine Einheit erhöht (alle anderen UVs bleiben konstant).

Gedacht in Log-Odds:

- a entspricht den Log-Odds von Personen (z.B. für Behandlungserfolg), die auf allen UVs den Wert 0 haben.
- β_j entspricht Zunahme der Log-Odds, falls sich die UV j um eine Einheit erhöht (alle anderen UVs bleiben konstant).

Multiple logistische Regression

Zurück zum Beispiel Beispiel:

- AV : Erfolg einer Gesprächstherapie (dummy-codiert)
 - $Y_i = 0$ falls nein
 - $Y_i = 1$ falls ja
- UV_1 : (stetige) gemessene Empathie der Therapeut:in
- UV_2 : (diskreter) erhobener Bindungsstil der Patient:in

Frage: Wie lässt sich der Zusammenhang zwischen den Variablen beschreiben?

Beispiel - nominalskalierte (diskrete) AV:

Behandlungserfolg	Empathie	Bindungsstil
ja	8	sicher
ja	9	sicher
nein	4	unsicher
ja	15	sicher
ja	10	sicher
nein	5	unsicher
nein	10	unsicher
nein	7	unsicher
ja	10	sicher
ja	15	sicher
nein	7	unsicher
nein	6	unsicher
nein	4	unsicher
nein	11	sicher
ja	11	sicher
nein	7	unsicher
ja	6	sicher

Multiple logistische Regression mit zentrierter UV

```
model = glm(Behandlungserfolg ~ center(Empathie) * Bindungsstil, data = data, family = "binomial")
car::Anova(model)

## Analysis of Deviance Table (Type II tests)
##
## Response: Behandlungserfolg
##              LR Chisq Df Pr(>Chisq)
## center(Empathie) 20.9213  1 0.000004786 ***
## Bindungsstil      5.0032  1   0.0253 *
## center(Empathie):Bindungsstil 1.6786  1   0.1951
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Omnibustest

- Haupteffekt Empathie ist signifikant ($\chi^2_1 = 20.92, p < .001$)
- Haupteffekt Bindungsstil ist signifikant ($\chi^2_1 = 5.00, p = .025$)
- Keine signifikante Interaktion ($\chi^2_1 = 1.68, p = .195$)

→ Empathie und Bindungsstil können Behandlungserfolg beide signifikant vorhersagen (Haupteffekte).

→ Empathie scheint bei sicher und unsicher gebundenen Personen gleich wichtig für Behandlungserfolg zu sein (Interaktion).

Logistische Regression

Effektstärken, Stichprobenplanung und Regressionsdiagnostik (s.h. Bühner et al. 2023)

- Auch für die logistische Regression spielen in der Praxis Effektstärken, Stichprobenplanung und Regressionsdiagnostik eine wichtige Rolle. Bei allen drei Themen existiert aber häufig keine eindeutig beste Variante, weshalb wir hier auf eine Beschreibung der Methoden verzichten.
- Bemerkung 1: In der logistischen Regression gibt es mehrere Möglichkeiten, eine globale Effektstärke in Anlehnung an R^2 in der linearen Regression zu definieren.
- Bemerkung 2: Für die Stichprobenplanung in der logistischen Regression müssen meist individuell Daten simuliert werden.
- Bemerkung 3: Für die Regressionsdiagnostik in der logistischen Regression gibt es mehrere Möglichkeiten, Residuen mit sinnvollen Eigenschaften zu definieren.

Take-aways

- Die **logistische Regression** ist ein statistisches Modell zur Vorhersage von binären oder kategorialen Ergebnissen (z.B. Ja/Nein, Erfolg/Misserfolg oder Krank/gesund).
- Mithilfe der **logistischen Funktion** (aka. sigmoide Funktion) wird Wertebereich zwischen 0 und 1 begrenzt (Wahrscheinlichkeiten).
- Modellparameter werden in **Log-Odds** angegeben (linear skaliert). Diese können durch Exponierung in besser interpretierbare **Odds** umgewandelt werden.
- **Statistische Inferenz** (Hypothesentesten) funktioniert entweder über Prüfung einzelner Modellparameter vs. 0 oder mittels Omnibustest (χ^2 -verteilt).
- Einfache logistische Regression kann zur **multiplen logistischen Regression** erweitert werden.
- **Powerberechnung** kann kompliziert sein und Simulation erfordern, **Effektstärken** werden uneinheitlich angegeben.