

Zusammenhang zwischen Alter, Geschlecht und Beziehungsstatus

Karsten Lübke

22.09.2018

Forschungsfrage

Der Zeitpunkt des Eintreten in verschiedene Lebensphasen (z. B. Verlassen des Elternhauses) variiert über die Kulturen und Jahre, siehe z. B. (Perelli-Harris and Amos 2015). Gilt dies auch für FOM Studierende? Gibt es Unterschiede im Beziehungsstatus zwischen Männern und Frauen und variiert dieser mit dem Alter, sind also z. B. ältere Studierende häufiger in einer festen Partnerschaft als jüngere Studierende?

Studiendesign

Freiwillige Online Umfrage bei FOM Studierenden unter der Leitung von Dipl.-Psych. Eva Wacker. D. h. es handelt sich um eine Gelegenheitsstichprobe.

Die relevanten Variablen sind manifest:

- Alter: numerisch, verhältnisskaliert, diskret erhoben.
- Geschlecht: kategorial nominal.
- Beziehungsstatus: kategorial nominal, binär. Antwort auf die Frage: "Sind Sie derzeit in einer festen Partnerschaft?"

Datenerhebung

Umfrage über SoSci Survey, vom 1.3.2018 bis 30.4.2018. Die Verbreitung erfolgte über einem Link im Online-Campus. Die Rohdaten liegen im **xlsx** (Excel) Format vor.

```
# Ggf. Paket readxl vorab installieren
# install.packages("readxl")

# Paket laden
library(readxl)

# Daten einlesen
Umfrage <- read_excel("Daten_SoSe2018_FOM_Wacker.xlsx")
```

Datenanalyse

Datenvorverarbeitung

Relevante Variablen auswählen:

```
# Ggfs. Paket mosaic installieren
# install.packages("mosaic")
```

```
# Paket laden
library(mosaic)
```

```
Umfrage <- Umfrage %>%
  select(Partnerschaft, Alter, Geschlecht)
```

In den Rohdaten liegen die Werte codiert vor, d. h. z. B. Die Antwort *Ja* auf die Frage “Sind Sie derzeit in einer festen Partnerschaft?” liegt als eine *1* vor. Dies wird zunächst umcodiert. Darüberhinaus werden die kategorialen Variablen als Faktoren in R definiert und evtl. fehlende Werte gelöscht.

```
Umfrage <- Umfrage %>%
  mutate(Partnerschaft = case_when(Partnerschaft == 1 ~ "Ja",
                                   Partnerschaft == 2 ~ "Nein")) %>%
  mutate(Geschlecht = case_when(Geschlecht == 1 ~ "weiblich",
                                Geschlecht == 2 ~ "männlich")) %>%
  mutate(Partnerschaft = factor(Partnerschaft, levels = c("Nein", "Ja"))) %>%
  mutate(Geschlecht = factor(Geschlecht)) %>%
  na.omit()
```

Erste Übersicht und Kontrolle der Daten:

```
inspect(Umfrage)
```

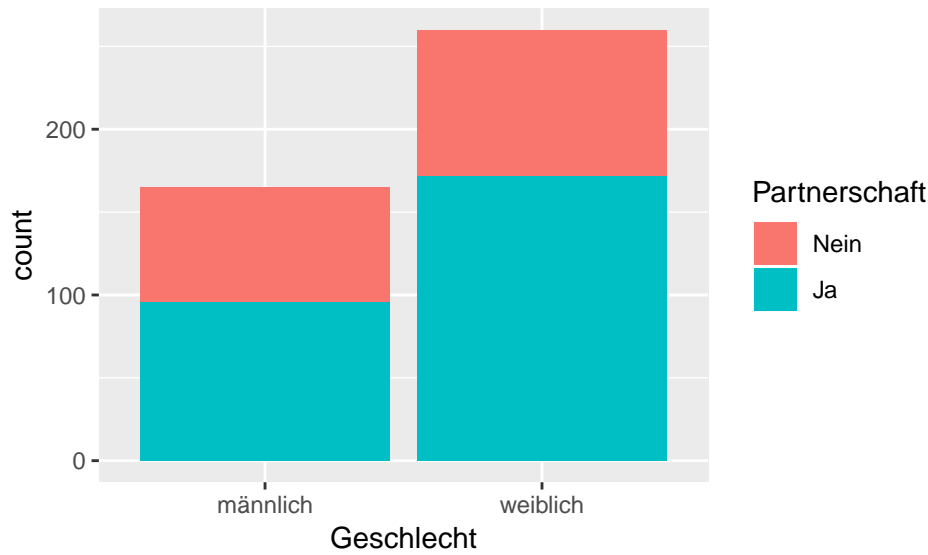
```
##
## categorical variables:
##      name  class levels  n missing
## 1 Partnerschaft factor      2 425      0
## 2   Geschlecht factor      2 425      0
##                                     distribution
## 1 Ja (63.1%), Nein (36.9%)
## 2 weiblich (61.2%), männlich (38.8%)
##
## quantitative variables:
##      name  class min Q1 median Q3 max  mean      sd  n missing
## 1 Alter numeric  18 22    24 27  39 25.00706 4.373747 425      0
```

Es liegen $n = 425$ Beobachtungen der drei Variablen vor.

Grafiken

Beziehungsstatus:

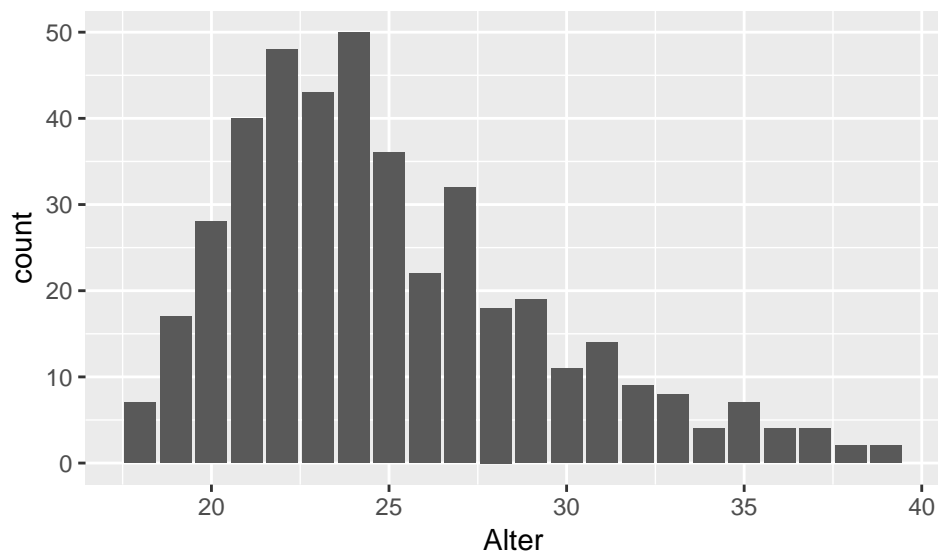
```
gf_bar(~ Geschlecht, fill=~Partnerschaft, data = Umfrage)
```



In der Stichprobe sind mehr Personen in einer Partnerschaft als nicht in einer Partnerschaft und insgesamt mehr Frauen als Männer. Frauen sind häufiger in einer festen Partnerschaft als Männer.

Alter:

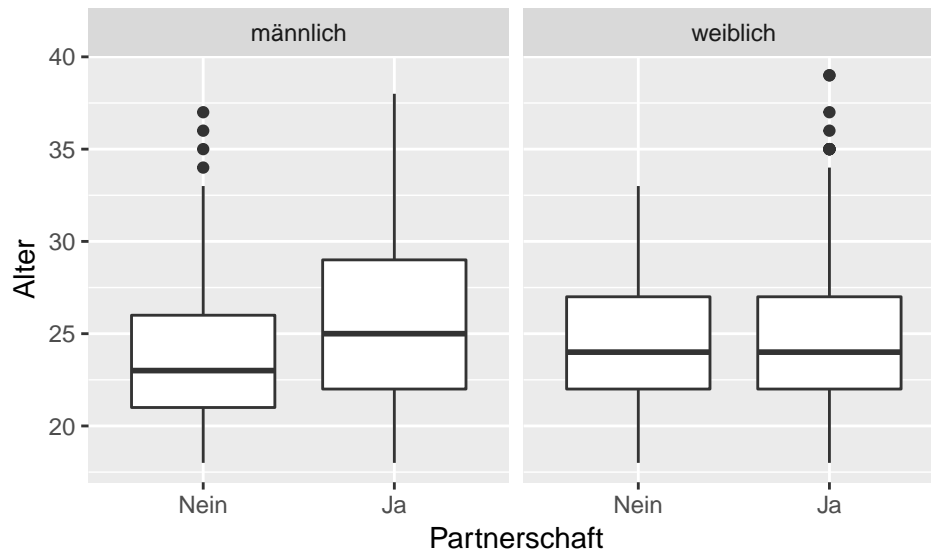
```
gf_bar( ~ Alter, data = Umfrage)
```



Das Alter ist insgesamt rechtsschief verteilt. Die jüngsten sind 18 Jahre, die ältesten knapp unter 40.

Alter je Beziehungsstatus und Geschlecht:

```
gf_boxplot(Alter ~ Partnerschaft | Geschlecht, data = Umfrage)
```



Untere- und Obere Quartil des Alters sind bei Personen die nicht in einer Partnerschaft leben bei den Männern leicht geringer, bei den Frauen ist kein Unterschied zu erkennen. Es gibt Ausreißer im Alter nach oben.

Kennzahlen

Die deskriptiven Kennzahlen bestätigen die visuellen Eindrücke:

Partnerschaft:

```
tally(Partnerschaft ~ Geschlecht, data = Umfrage, format = "proportion")
```

```
##           Geschlecht
## Partnerschaft  männlich weiblich
##           Nein 0.4181818 0.3384615
##           Ja  0.5818182 0.6615385
```

Während 66% der Frauen in einer festen Partnerschaft sind, sind es 58% der Männer.

Alter:

```
favstats(Alter ~ Partnerschaft, data = Umfrage)
```

```
##   Partnerschaft min Q1 median Q3 max   mean    sd  n missing
## 1           Nein  18 21    24 27  37 24.35032 4.023754 157      0
## 2            Ja  18 22    24 28  39 25.39179 4.529496 268      0
```

```
favstats(Alter ~ Geschlecht, data = Umfrage)
```

```
##   Geschlecht min Q1 median Q3 max   mean    sd  n missing
## 1   männlich  18 22    24 28  38 25.18182 4.798720 165      0
## 2   weiblich  18 22    24 27  39 24.89615 4.086987 260      0
```

```
favstats(Alter ~ Geschlecht + Partnerschaft, data = Umfrage)
```

```
##   Geschlecht.Partnerschaft min Q1 median Q3 max   mean    sd  n
## 1   männlich.Nein         18 21    23 26  37 24.08696 4.539821 69
## 2   weiblich.Nein         18 22    24 27  33 24.55682 3.581106 88
## 3   männlich.Ja          18 22    25 29  38 25.96875 4.848121 96
## 4   weiblich.Ja          18 22    24 27  39 25.06977 4.322634 172
##   missing
```

```
## 1      0
## 2      0
## 3      0
## 4      0
```

Im Mittelwert sind diejenigen, die in einer festen Partnerschaft leben ein Jahr älter, als diejenigen, die nicht in einer festen Partnerschaft leben (25.4 Jahre statt 24.4 Jahre). Auch sind die männlichen Studierenden im Mittelwert leicht älter als die weiblichen (25.2 zu 24.9).

Modellierung/ Inferenz

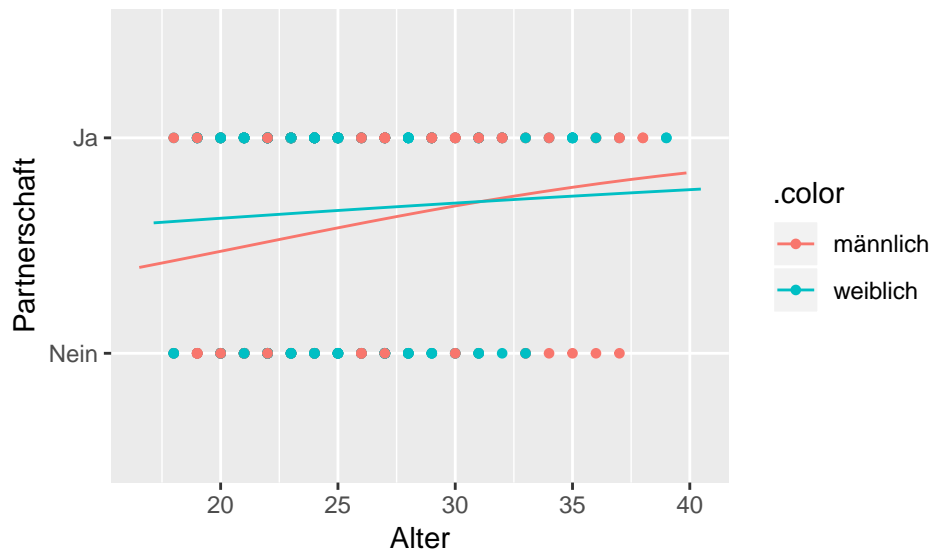
Zur Modellierung der Wahrscheinlichkeit in einer festen Partnerschaft zu leben wird ein logistisches Modell inkl. Wechselwirkung verwendet:

$$P(\text{Partnerschaft=Ja}) = \frac{e^{\beta_0 + \beta_{\text{Alter}} \cdot \text{Alter} + \beta_{\text{Geschlecht}} \cdot \text{Geschlecht} + \beta_{\text{Alter:Geschlecht}} \cdot (\text{Alter:Geschlecht})}}{1 + e^{\beta_0 + \beta_{\text{Alter}} \cdot \text{Alter} + \beta_{\text{Geschlecht}} \cdot \text{Geschlecht} + \beta_{\text{Alter:Geschlecht}} \cdot (\text{Alter:Geschlecht})}}$$

Wobei hier gilt:¹

$$\text{Geschlecht} = \begin{cases} 1, & \text{weiblich} \\ 0, & \text{nicht weiblich} \end{cases}$$

```
ergglm <- glm(Partnerschaft ~ Alter*Geschlecht, data = Umfrage, family = "binomial")
plotModel(ergglm)
```



```
summary(ergglm)
```

```
##
## Call:
## glm(formula = Partnerschaft ~ Alter * Geschlecht, family = "binomial",
##      data = Umfrage)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7945  -1.3603   0.8720   0.9424   1.2996
```

¹ siehe levels(Umfrage\$Geschlecht)

```
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.86467    0.90396  -2.063   0.0391 *
## Alter           0.08789    0.03596   2.444   0.0145 *
## Geschlechtweiblich 1.75162    1.22428   1.431   0.1525
## Alter:Geschlechtweiblich -0.05632    0.04879  -1.154   0.2484
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 559.85  on 424  degrees of freedom
## Residual deviance: 549.72  on 421  degrees of freedom
## AIC: 557.72
##
## Number of Fisher Scoring iterations: 4
```

Da $\hat{\beta}_{Alter} = 0.09 > 0$ steigt im Modell in der Stichprobe mit zunehmenden Alter die Wahrscheinlichkeit in einer festen Partnerschaft zu leben. Die Nullhypothese es gibt keinen Zusammenhang zwischen Alter und Partnerschaft (Ja/Nein) ($H_0 : \beta = 0$) wird mit einem p-Wert von 0.0145 zum Niveau $\alpha = 0.05$ verworfen. Die Wahrscheinlichkeit in einer festen Partnerschaft zu leben ist in diesem Modell für Frauen höher als für Männer ($\hat{\beta}_{Geschlecht} = 1.75 > 0$), allerdings ist dieser beobachtete Effekt mit einem p-Wert von 0.15 nicht besonders unwahrscheinlich, wenn gilt $\beta_{Geschlecht} = 0$. Der Anstieg der Wahrscheinlichkeit für eine feste Partnerschaft ist mit zunehmenden Alter für Frauen zwar geringer als für die Männer ($\hat{\beta}_{Alter:Geschlecht} = -0.06 < 0$), aber diese beobachtete Wechselwirkung ist ebenfalls nicht besonders unwahrscheinlich, wenn wir annehmen $\beta_{Alter:Geschlecht} = 0$.

Dieses Ergebnis wird durch Bootstrapping bestätigt:

```
set.seed(1896) # Reproduzierbarkeit
Bootvtlg <- do(10000) * glm(Partnerschaft ~ Alter*Geschlecht,
                           data = resample(Umfrage), family = "binomial")
confint(Bootvtlg)
```

```
##              name      lower      upper level    method
## 1      Intercept -3.89239670 -0.12410743  0.95 percentile
## 2          Alter  0.01967856  0.17170203  0.95 percentile
## 3  Geschlechtweiblich -0.59207564  4.31216548  0.95 percentile
## 4 Alter.Geschlechtweiblich -0.15936617  0.03726076  0.95 percentile
##      estimate
## 1 -1.86467426
## 2  0.08788651
## 3  1.75162398
## 4 -0.05631593
```

Nur der Koeffizient für die Variable Alter überdeckt nicht die 0.

Die Chance² in einer festen Partnerschaft zu leben ändert sich in der Stichprobe und dem Modell mit jedem Jahr bei den Männern um den Faktor $e^{\hat{\beta}_{Alter}} = e^{0.09} = 1.09$, bei den Frauen um $e^{\hat{\beta}_{Alter} + \hat{\beta}_{Alter:Geschlecht}} = e^{0.09 - 0.06} = 1.03$.

Wird für diejenigen, für die die geschätzte Wahrscheinlichkeit einer Partnerschaft größer als 0.5 ist, also $\hat{P}(\text{Partnerschaft}=\text{Ja}) > 0.5$ vorausgesetzt, dass sie in einer Partnerschaft leben,

$$^2Odds = \frac{P(\text{Partnerschaft}=\text{Ja})}{1 - P(\text{Partnerschaft}=\text{Ja})}$$

```
Umfrage <- Umfrage %>%
  mutate(pdach_Partnerschaft = predict(ergglm, type = "response")) %>%
  mutate(PartnerschaftDach = case_when(pdach_Partnerschaft <= 0.5 ~ "Nein",
                                       pdach_Partnerschaft > 0.5 ~ "Ja"))
```

so ergibt sich eine Fehlerrate auf den Datensatz von:

```
prop( ~ (Partnerschaft!=PartnerschaftDach), data=Umfrage)
```

```
## prop_TRUE
## 0.3529412
```

Dies ist nur wenig besser als wenn für alle der Modalwert “Ja” vorausgesagt wird:

```
prop( ~ (Partnerschaft!="Ja"), data=Umfrage)
```

```
## prop_TRUE
## 0.3694118
```

Schlussfolgerungen

Innerhalb dieser Stichprobe von FOM Studierenden, ohne Berücksichtigung von Kovariablen (Attraktivität etc.) gibt es innerhalb einer logistischen Regression einen Zusammenhang zwischen dem Alter und in fester Partnerschaft lebend: mit zunehmenden Alter steigt im Modell der Stichprobe die Wahrscheinlichkeit in einer festen Partnerschaft zu leben. Bei Frauen ist dieser Zusammenhang von der Größe her schwächer und Frauen haben insgesamt eine höhere Wahrscheinlichkeit in einer festen Partnerschaft zu leben, aber diese Unterschiede sind nicht unwahrscheinlich unter $H_0 : \beta_j = 0$. Allerdings ist die Modellgüte mit einer Fehlklassifikationsrate von 0.35 nicht besonders gut.

Aufgrund der Gelegenheitsstichprobe unter FOM Studierenden ist die Übertragbarkeit (externe Validität) eingeschränkt, auch ist aufgrund der Beobachtungsstudie kein Kausalschluss (interne Validität) möglich.

Darüberhinaus stellt sich die Frage ob alle das gleiche unter einer “festen Partnerschaft” verstehen und ob wahrheitsgemäß geantwortet wurde.

Danksagung

Die Daten wurden von Frau Eva Wacker erhoben und zur Verfügung gestellt.

Anhang: Versionshinweise

- Datum erstellt: 2019-06-12
- R Version: 3.6.0

Verwendete Pakete:

- mosaic Version: 1.5.0

Literatur

Perelli-Harris, Brienna, and Mark Amos. 2015. "Changes in partnership patterns across the life course: An examination of 14 countries in Europe and the United States." *Demographic Research* 33 (6): 145–78. <https://doi.org/10.4054/DemRes.2015.33.6>.