

Lösungen zu den Aufgaben

1. Aufgabe

Laden Sie den Datensatz `toothgrowth`, da wird ein einfaches Experiment geschildert:

Forschungsfrage:

Hat die Darreichungsform (supp, OJ vs. VC) einen (kausalen) Effekt auf die AV?

Aufgabe

Berechnen Sie ein 95%-HDI oder 95%-PI für den Effekt!

Lösung

1. Schritt: Modell berechnen

Zur Erinnerung: Bei der Regressionsformel gilt immer $av \sim uv$.

2. Schritt: Posteriori-Verteilung betrachten

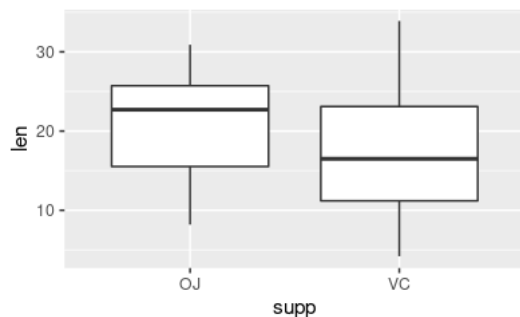
Mit `summary()` kriegt man einen guten Überblick:

```
##
## Model Info:
## function:      stan_glm
## family:        gaussian [identity]
## formula:       len ~ supp
## algorithm:     sampling
## sample:        4000 (posterior sample size)
## priors:        see help('prior_summary')
## observations:  60
## predictors:    2
##
## Estimates:
##           mean    sd   10%   50%   90%
## (Intercept) 20.7    1.4  18.8  20.7  22.5
## suppVC      -3.7    2.0   -6.2  -3.7  -1.1
## sigma        7.6    0.7    6.7   7.5   8.5
##
## Fit Diagnostics:
##           mean    sd   10%   50%   90%
## mean_PPD 18.8    1.4  17.1  18.8  20.6
##
## The mean_ppd is the sample average posterior predictive distribution of the outcome variable (for details see help('summary.stanreg'))
##
## MCMC diagnostics
##           mcse  Rhat  n_eff
## (Intercept)  0.0   1.0  3541
## suppVC       0.0   1.0  3653
## sigma        0.0   1.0  3465
## mean_PPD     0.0   1.0  3872
## log-posterior 0.0   1.0  1827
##
## For each parameter, mcse is Monte Carlo standard error, n_eff is a crude measure of effective sample size, and Rhat is the potential s
```

Aber man kann z.B. auch mit `posterior_interval()` sich sein Wunsch-Intervall ausgeben lassen:

```
##           2.5%    97.5%
## (Intercept) 17.847889 23.4781348
## suppVC      -7.621699  0.1361652
## sigma        6.326847  9.1958483
```

Schauen wir uns mal zum Vergleich die Stichproben-Daten an:



2. Aufgabe

Laden Sie den Datensatz `toothgrowth`, da wird ein einfaches Experiment geschildert:

Forschungsfrage:

Hat die Dosis (`dose`) einen (kausalen) Effekt auf die AV?

Wir gehen mal einfach davon aus, dass der Faktor experimentell (also randomisiert und auf Störeffekte hin kontrolliert) verabreicht wurde. Sonst wäre eine Kausalinterpretation nicht (ohne Weiteres) möglich.

Aufgabe

Berechnen Sie ein 95%-HDI oder 95%-PI für den Effekt!

Lösung

1. Schritt: Modell berechnen

Zur Erinnerung: Bei der Regressionsformel gilt immer $av \sim uv$.

2. Schritt: Posteriori-Verteilung betrachten

Mit `summary()` kriegt man einen guten Überblick:

```
##
## Model Info:
## function:      stan_glm
## family:        gaussian [identity]
## formula:       len ~ dose
## algorithm:     sampling
## sample:        4000 (posterior sample size)
## priors:        see help('prior_summary')
## observations:  60
## predictors:    2
##
## Estimates:
##           mean    sd   10%   50%   90%
## (Intercept)  7.4    1.3   5.9   7.4   9.1
## dose         9.8    1.0   8.5   9.8  11.0
## sigma        4.7    0.5   4.1   4.7   5.3
##
## Fit Diagnostics:
##           mean    sd   10%   50%   90%
## mean_PPD 18.8    0.9 17.7  18.8  19.9
##
## The mean_ppd is the sample average posterior predictive distribution of the outcome variable (for details see help('summary.stanreg'))
##
## MCMC diagnostics
##           mcse Rhat n_eff
## (Intercept)  0.0  1.0  3441
## dose         0.0  1.0  3528
## sigma        0.0  1.0  2831
## mean_PPD     0.0  1.0  3610
## log-posterior 0.0  1.0 1694
##
## For each parameter, mcse is Monte Carlo standard error, n_eff is a crude measure of effective sample size, and Rhat is the potential
```

Aber man kann z.B. auch mit `posterior_interval()` sich sein Wunsch-Intervall (PI) ausgeben lassen: Darreichungsform

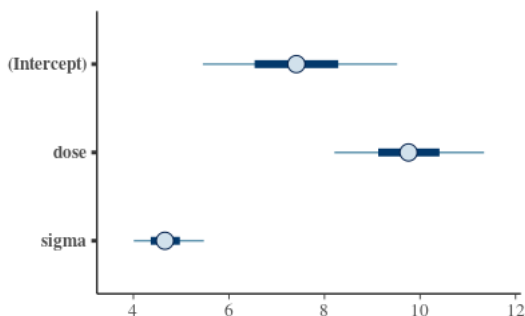
```
##           2.5%    97.5%
## (Intercept) 5.005455  9.961171
## dose       7.857502 11.599158
## sigma      3.905292  5.682683
```

Das Modell zeigt einen positiven Effekt für `dose`:

Pro Einheit von `dose` steigt die Zahnlänge (`len`) um ca. 8-12 mm im Schnitt (laut unserem Modell).

Null ist nicht im Intervall enthalten; die Nullhypothese ist demnach auszuschließen (falls das jemanden interessiert).

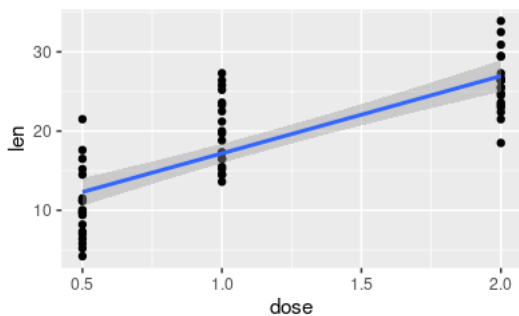
Das können wir auch plotten:



Man kann sich auch ein HDI ausgeben, aber nicht mit `rstanarm`, das kann nur PI. Aber mit `{bayesTestR}` geht das:

```
## Highest Density Interval
##
## Parameter |          95% HDI
## -----
## (Intercept) | [5.12, 10.04]
## dose       | [7.85, 11.58]
```

Schauen wir uns mal zum Vergleich die Stichproben-Daten an:



Man sieht deutlich einen positiven Effekt: Die Regressionsgerade steigt.

3. Aufgabe

Laden Sie den Datensatz `toothgrowth`, da wird ein einfaches Experiment geschildert:

Forschungsfragen:

1. Hat die Dosis (`dose`) einen (kausalen) Effekt auf die AV?
2. Hat die Darreichungsform (`supp`) einen (kausalen) Effekt auf die AV?
3. Gibt es einen Interaktionseffekt zwischen beiden UV?

Wir sind also nicht nur an dem Interaktionseffekt interessiert, sondern auch an den beiden Haupteffekten (also die Effekte der beiden UV). Insgesamt haben wir also drei Einflussgrößen in unserem Modell.

Wir gehen mal einfach davon aus, dass der Faktor experimentell (also randomisiert und auf Störeffekte hin kontrolliert) verabreicht wurde. Sonst wäre eine Kausalinterpretation nicht (ohne Weiteres) möglich.

Aufgabe

Berechnen Sie ein 95%-HDI oder 95%-PI für die Effekte!

Lösung

1. Schritt: Modell berechnen

Zur Erinnerung: Bei der Regressionsformel gilt immer $av \sim uv$.

2. Schritt: Posteriori-Verteilung betrachten

Mit `summary()` kriegt man einen guten Überblick:

```
##
## Model Info:
## function:      stan_glm
## family:        gaussian [identity]
## formula:       len ~ dose + supp + dose:supp
## algorithm:     sampling
## sample:        4000 (posterior sample size)
## priors:        see help('prior_summary')
## observations:  60
## predictors:    4
##
## Estimates:
##           mean    sd   10%   50%   90%
## (Intercept) 11.6   1.6    9.5  11.6  13.6
## dose         7.8   1.2    6.3   7.8   9.4
## suppVC      -8.2   2.3  -11.1  -8.2  -5.4
## dose:suppVC   3.9   1.7    1.7   3.9   6.1
## sigma        4.1   0.4    3.7   4.1   4.7
##
## Fit Diagnostics:
##           mean    sd   10%   50%   90%
## mean_PPD 18.8    0.8  17.9  18.8  19.8
##
## The mean_ppd is the sample average posterior predictive distribution of the outcome variable (for details see help('summary.stanreg'))
```

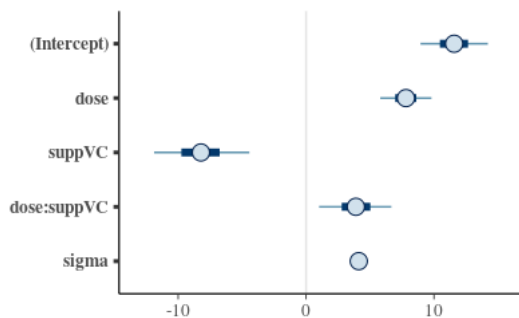
```
##
## MCMC diagnostics
##               mcse Rhat n_eff
## (Intercept)   0.0  1.0  2018
## dose          0.0  1.0  1981
## suppVC        0.1  1.0  1760
## dose:suppVC    0.0  1.0  1739
## sigma         0.0  1.0  2816
## mean_PPD      0.0  1.0  3573
## log-posterior 0.0  1.0  1694
##
## For each parameter, mcse is Monte Carlo standard error, n_eff is a crude measure of effective sample size, and Rhat is the potential
```

Aber man kann z.B. auch mit `posterior_interval()` sich sein Wunsch-Intervall (PI) ausgeben lassen:

```
##               2.5%   97.5%
## (Intercept)  8.4323454 14.699082
## dose        5.3930939 10.167450
## suppVC     -12.7465869 -3.755986
## dose:suppVC  0.4769072  7.241870
## sigma       3.4638635  5.007826
```

Alle drei Modelle zeigen Schätzbereiche an, in denen die Null nicht enthalten ist; die Nullhypothese ist demnach jeweils auszuschließen (falls das jemanden interessiert).

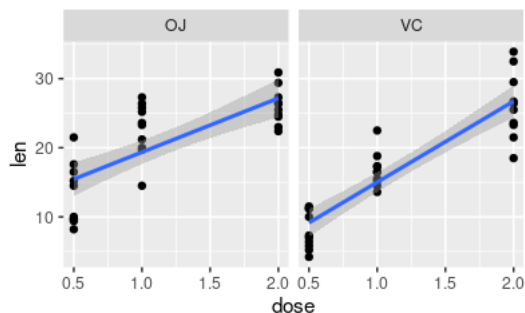
Das können wir auch plotten:



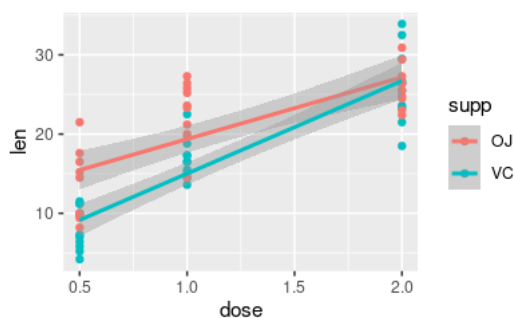
Man kann sich auch ein HDI ausgeben, aber nicht mit `rstanarm`, das kann nur PI. Aber mit `{bayesTestR}` geht das:

```
## Highest Density Interval
##
## Parameter |          95% HDI
## -----|-----
## (Intercept) | [ 8.50, 14.74]
## dose       | [ 5.47, 10.22]
## suppVC     | [-12.84, -3.88]
## dose:suppVC | [ 0.55, 7.27]
```

„ Schauen wir uns mal zum Vergleich die Stichproben-Daten an:



Oder so, vielleicht besser:



Man sieht deutlich einen positiven Effekt für *dose*: Die Regressionsgerade steigt. Außerdem gibt es - in den Stichprobendaten zumindest - einen Interaktionseffekt: Die Geraden sind nicht parallel. Drittens ist der Mittelwert der Gruppe *oj* augenscheinlich höher als in der anderen Gruppe: Die Darbietungsform hat offenbar einen Effekt.

4. Aufgabe

Das Testen von Nullhypothesen wird u.a. deswegen kritisiert, weil die Nullhypothese zumeist apriori als falsch bekannt ist, weswegen es keinen Sinne mache, so die Kritiker, sie zu testen.

Nennen Sie ein Verfahren von John Kruschke, das einen Äquivalenzbereich testet und insofern eine Alternative zum Testen von Nullhypothesen anbietet.

Hinweise:

- Geben Sie nur Kleinbuchstaben ein.
- Geben Sie nur ein einziges Wort ein.

Lösung

rope

5. Aufgabe

Im Datensatz *mtcars*: Ist der (mittlere) Unterschied im Spritverbrauch zwischen den beiden Gruppen *Automatik* vs. *Schaltgetriebe* vernachlässigbar?

Definieren Sie selber, was "vernachlässigbar klein" bedeutet. Oder greifen Sie auf die Definition "höchstens eine Meile" zurück.

Prüfen Sie rechnerisch, anhand des angegebenen Datensatzes, folgende Behauptung:

Behauptung: "Der Unterschied ist vernachlässigbar klein!"

Wählen Sie die Antwortoption, die am besten zu der obigen Behauptung passt!

Hinweise:

- Sie benötigen einen Computer, um diese Aufgabe zu lösen.
- Verwenden Sie die statistischen Methoden, die im Unterricht behandelt wurden.
- Verwenden Sie Ansätze aus der Bayes-Statistik zur Lösung dieser Aufgabe.

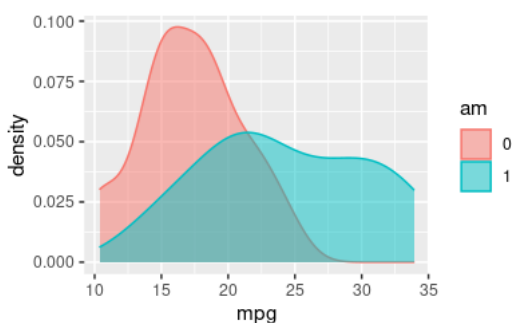
Antwortoptionen:

- Ja, die Behauptung ist korrekt.
- Nein, die Behauptung ist falsch.
- Die Daten sind bzw. das Modell nicht konkludent; es ist keine Entscheidung über die Behauptung möglich.
- Auf Basis der bereitgestellten Informationen ist keine Entscheidung möglich über die Behauptung.

Lösung

Zur ersten Orientierung erstellen wir uns, rein deskriptiv, eine Darstellung des Spritverbrauchs beider Gruppen, z.B. so:

```
mtcars %>%  
  mutate(am = factor(am)) %>%  
  ggplot() +  
    aes(x = mpg, color = am, fill = am) +  
    geom_density(alpha = .5)
```



Man sieht direkt, dass es substantielle Unterschiede zwischen den beiden Gruppen gibt. Vermutlich wird das Modell, das wir gleich berechnen, uns wenig überraschen, sondern den deskriptiven Befund widerspiegeln.

Modell berechnen:

```
library(rstanarm)
library(tidyverse)
data(mtcars)

m1_mtcars <- stan_glm(mpg ~ am, data = mtcars, refresh = 0)
```

Posteriori-Verteilung betrachten:

```
m1_mtcars

## stan_glm
## family:      gaussian [identity]
## formula:      mpg ~ am
## observations: 32
## predictors:    2
## -----
##              Median MAD_SD
## (Intercept)  17.2    1.1
## am           7.2    1.7
##
## Auxiliary parameter(s):
##              Median MAD_SD
## sigma 5.0    0.6
## -----
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg

coef(m1_mtcars)

## (Intercept)      am
##  17.155545    7.239845

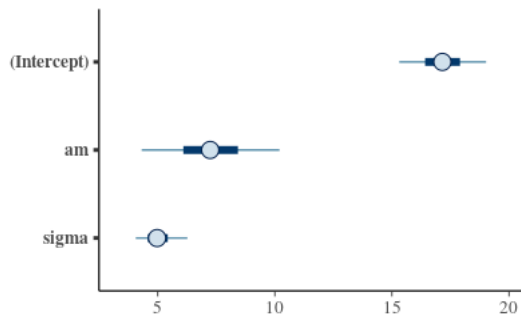
posterior_interval(m1_mtcars, prob = .95)

##              2.5%    97.5%
## (Intercept) 14.901391 19.445887
## am          3.640913 10.771611
## sigma       3.915098  6.511744
```

Spuckt ein PI aus, kein HDI (HDI noch nicht implementiert in rstanarm).

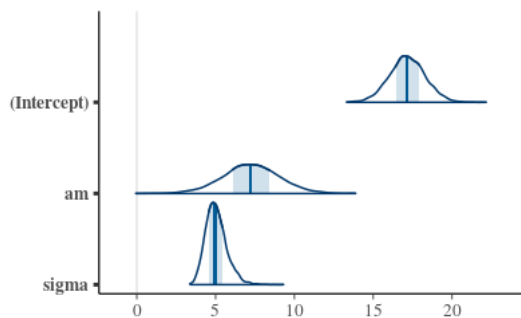
Visualisieren der Posteriori-Verteilung:

```
plot(m1_mtcars)
```



Oder als Histogramm:

```
library(bayesplot)
mcmc_areas(m1_mtcars)
```



Man sieht direkt, dass der Unterschied komplett außerhalb des Rope liegt.

Rope berechnen:

```
library(bayestestR)
rope_m1 <- rope(m1_mtcars, range = c(-1, 1)) # ±1 Meile Unterschied
```

Rope visualisieren:

```
plot(rope_m1)
```

```
## Error: Failed at retrieving data :( Please provide original model or data through the `data` argument
```

Man sieht, dass der "Berg" - die Posteriori-Verteilung bzw. der Bereich plausibler Werte - außerhalb des Rope-Bereichs liegt.

Wir können also die Hypothese, dass der Unterschied zwischen beiden Gruppen praktisch Null ist, verwerfen.

Natürlich ist das nur ein deskriptiver Befund, wir können nichts dazu sagen, ob der Unterschied auch ein kausaler Effekt ist.

Alternative Rope-Definition: Z-Standardisieren.

Ein kleiner Effekt ist, laut Kruschke 2018, ein Unterschied der nicht größer ist als ± 0.1 SD.

```
m2_mtcars <-  
  mtcars %>%  
  mutate(mpg_z = scale(mpg)) %>%  
  stan_glm(mpg_z ~ am, data = ., refresh = 0)  
  
rope(m2_mtcars)  
  
## # Proportion of samples inside the ROPE [-0.10, 0.10]:  
##  
## Parameter | inside ROPE  
## -----  
## (Intercept) | 0.00 %  
## am | 0.00 %  
  
plot(rope(m2_mtcars))  
  
## Error: Failed at retrieving data :( Please provide original model or data through the `data` argument
```

- a. Falsch
- b. Richtig
- c. Falsch
- d. Falsch

6. Aufgabe

Einer der (bisher) größten Studien der Untersuchung psychologischer Konsequenzen (oder Korrelate) der Covid-Zeit ist die Studie [COVIDiStress](#).

Im Folgenden sollen Sie folgende Forschungsfrage untersuchen:

Ist der Zusammenhang von Stress (PSS10_avg, AV) und Neurotizismus (neu, UV) vernachlässigbar klein?

Den Datensatz können Sie so herunterladen (Achtung, groß):

```
osf_d_path <- "https://osf.io/cjxua/?action=download"  
  
d <- read_csv(osf_d_path)  
  
## Warning: One or more parsing issues, see `problems()` for details
```

Hinweise:

- Sie benötigen einen Computer, um diese Aufgabe zu lösen.
- Verwenden Sie die statistischen Methoden, die im Unterricht behandelt wurden.
- Verwenden Sie Ansätze aus der Bayes-Statistik zur Lösung dieser Aufgabe.
- Bei der Variable für Geschlecht können Sie sich auf Fälle begrenzen, die Männer und Frauen umfassen.
- Wandeln Sie die Variable für Geschlecht in eine binäre Variable - also Werte mit 0 und 1 - um.
- Alle Daten (und weitere Informationen) zum Projekt sind [hier](#) abgelegt.
- Eine Beschreibung der Variablen der Studie finden Sie [hier](#).
- Das Codebook findet sich [hier](#).

Antwortoptionen

- a. Ja
- b. Nein
- c. Die Daten sind nicht konkludent; es ist keine Entscheidung möglich.
- d. Auf Basis der bereitgestellten Informationen ist keine Entscheidung möglich.

Lösung

Pakete laden:

```
library(tidyverse)  
library(rstanarm)  
library(rstatix)  
library(bayestestR)
```

Wie groß ist der Datensatz (im Speicher) eigentlich, in Megabyte?

```
object.size(d) / 1024 / 1024  
  
## 156.8 bytes
```

Relevante Spalten auswählen:

```
d2 <-  
d %>%  
  select(PSS10_avg, neu)
```

Datensatz aufbereiten:

```
d3 <-  
d2 %>%  
  drop_na()
```

Modell berechnen:

```
m1 <-  
  stan_glm(PSS10_avg ~ neu,  
    refresh = 0,  
    data = d3)
```

Modellkoeffizienten auslesen:

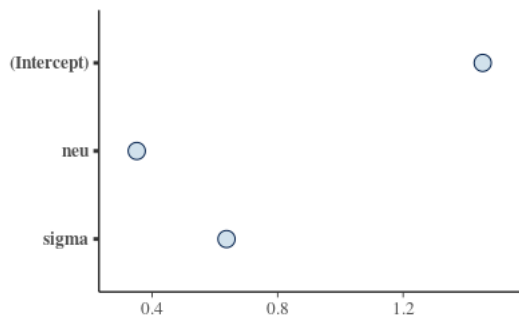
```
coef(m1)  
  
## (Intercept)      neu  
##  1.4535335  0.3509463
```

Posteriori-Verteilung auslesen:

```
posterior_interval(m1, prob = .89)  
  
##              5.5%      94.5%  
## (Intercept) 1.4430049 1.4640590  
## neu         0.3479444 0.3539071  
## sigma       0.6348510 0.6391500
```

Posteriori-Verteilung plotten:

```
plot(m1)
```



Rope berechnen:

```
rope_m1 <- rope(m1)
```

Rope visualisieren:

```
plot(rope_m1)
```

```
## Error: Failed at retrieving data :( Please provide original model or data through the `data` argument
```

- a. Falsch
- b. Wahr
- c. Falsch
- d. Falsch

7. Aufgabe

Einer der (bisher) größten Studien der Untersuchung psychologischer Konsequenzen (oder Korrelate) der Covid-Zeit ist die Studie [COVIDiStress](https://osf.io/cjxua/).

Im Folgenden sollen Sie folgende Forschungsfrage untersuchen:

Forschungsfrage:

Ist der Unterschied zwischen Männern und Frauen (Dem_gender) im Hinblick zum Zusammenhang von Stress (PSS10_avg, AV) und Neurotizismus (neu, UV) vernachlässigbar klein?

Den Datensatz können Sie so herunterladen (Achtung, groß):

```
osf_d_path <- "https://osf.io/cjxua/?action=download"  
d <- read_csv(osf_d_path)  
  
## Warning: One or more parsing issues, see `problems()` for details
```


Hinweise:

- Sie benötigen einen Computer, um diese Aufgabe zu lösen.
- Verwenden Sie die statistischen Methoden, die im Unterricht behandelt wurden.
- Verwenden Sie Ansätze aus der Bayes-Statistik zur Lösung dieser Aufgabe.
- Bei der Variable für Geschlecht können Sie sich auf Fälle begrenzen, die Männer und Frauen umfassen.
- Wandeln Sie die Variable für Geschlecht in eine binäre Variable - also Werte mit 0 und 1 - um.
- Alle Daten (und weitere Informationen) zum Projekt sind [hier](#) abgelegt.
- Eine Beschreibung der Variablen der Studie finden Sie [hier](#).

Antwortoptionen:

- a. Ja
- b. Nein
- c. Die Daten sind nicht konkludent; es ist keine Entscheidung möglich.
- d. Auf Basis der bereitgestellten Informationen ist keine Entscheidung möglich.

Lösung

Pakete laden:

```
library(tidyverse)
library(rstanarm)
library(rstatix)
library(bayestestR)
```

Relevante Spalten auswählen:

```
d2 <-
  d %>%
  select(PSS10_avg, neu, Dem_gender)
```

Das sind die Variablen:

- Stress
- Neurotizismus
- Geschlecht

Deskriptive Statistiken zum Datensatz:

```
d2 %>%
  get_summary_stats(type = "robust")

## # A tibble: 2 × 4
##   variable      n median   iqr
##   <chr>      <dbl> <dbl> <dbl>
## 1 neu      108367   3.33  1.33
## 2 PSS10_avg 116097   2.6   1
```

```
d2 %>%
  count(Dem_gender)

## # A tibble: 4 × 2
##   Dem_gender      n
##   <chr>      <int>
## 1 Female    90400
## 2 Male     33126
## 3 Other/would rather not say 1474
## 4 <NA>      306
```

Datensatz aufbereiten:

```
d3 <-
  d2 %>%
  filter(Dem_gender %in% c("Female", "Male")) %>%
  drop_na() %>%
  mutate(Female = ifelse(Dem_gender == "Female", 1, 0)) %>%
  select(-Dem_gender)
```

Check:

```
d3 %>%
  count(Female)

## # A tibble: 2 × 2
##   Female      n
##   <dbl> <int>
## 1      0 28371
## 2      1 78472
```

Check:

```
d3 %>%
  get_summary_stats()

## # A tibble: 3 × 13
##   variable      n   min   max median    q1    q3   iqr   mad   mean    sd    se
##   <chr>      <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 Female    106843     0     1     1     0     1     1     0    0.734 0.442 0.001
```

```
## 2 neu      106843      1      6  3.33 2.67  4      1.33 0.988 3.34  1.05 0.003
## 3 PSS10_avg 106843      1      5  2.6  2.1  3.1  1      0.741 2.62  0.735 0.002
## # ... with 1 more variable: ci <dbl>
```

Modell berechnen:

```
m1 <-
  stan_glm(PSS10_avg ~ neu + Female + PSS10_avg:Female,
    refresh = 0,
    data = d3)
```

Modellkoeffizienten auslesen:

```
coef(m1)

##      (Intercept)          neu      Female PSS10_avg:Female
##      2.1204064      0.1115501     -2.2920919      0.9222247
```

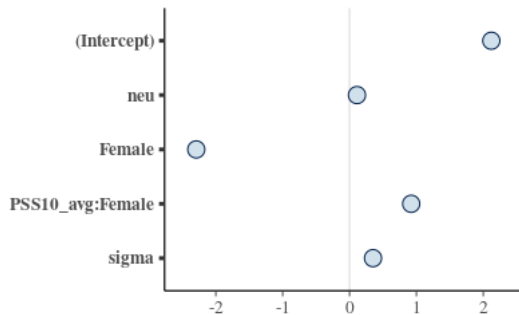
Posteriori-Verteilung auslesen:

```
posterior_interval(m1, prob = .89)

##              5.5%      94.5%
## (Intercept)  2.1139764  2.1268407
## neu          0.1097698  0.1133717
## Female      -2.3011271 -2.2835684
## PSS10_avg:Female 0.9193708  0.9253449
## sigma        0.3499503  0.3523743
```

Posteriori-Verteilung plotten:

```
plot(m1)
```



Rope berechnen:

```
rope(m1)

## # Proportion of samples inside the ROPE [-0.07, 0.07]:
##
## Parameter | inside ROPE
## -----
## (Intercept) | 0.00 %
## neu | 0.00 %
## Female | 0.00 %
## PSS10_avg:Female | 0.00 %
```

Rope visualisieren:

```
plot(rope(m1))
```

```
## Error: Failed at retrieving data :( Please provide original model or data through the `data` argument
```

- Falsch
- Wahr
- Falsch
- Falsch