Ch5

In this notebook, you will find my solutions to some exercises from Chapter 5 of *Statistical Rethinking* and the assigned exercises from this course.

Chapter 5

Book Exercises

Exercise 5M4

Data preparation. Standardize LDS members per 100 000 population. Merging with `WaffleDivorce` data.

```
library(dplyr)
```

```
Attache Paket: 'dplyr'
```

```
Die folgenden Objekte sind maskiert von 'package:stats':

    filter, lag
```

```
Die folgenden Objekte sind maskiert von 'package:base':

    intersect, setdiff, setequal, union
```

```
library(rethinking)
```

```
Lade nötiges Paket: cmdstanr
```

```
This is cmdstanr version 0.8.0
```

```
- CmdStanR documentation and vignettes: mc-stan.org/cmdstanr
```

```
- CmdStan path: /Users/eleonora/Documents/PhD/Statistics/cmdstan
```

```
- CmdStan version: 2.36.0
```

```
Lade nötiges Paket: posterior
```

```
This is posterior version 1.6.1
```

```

Attache Paket: 'posterior'
```

```
Die folgenden Objekte sind maskiert von 'package:stats':

    mad, sd, var
```

```
Die folgenden Objekte sind maskiert von 'package:base':

    %in%, match
```

```
Lade nötiges Paket: parallel
```

```
rethinking (Version 2.42)
```

```

Attache Paket: 'rethinking'
```

```
Das folgende Objekt ist maskiert 'package:stats':

    rstudent
```

```r
# upload data fro csv
lds <- read.csv("../data/lds.csv") %>%
  mutate(lds_prop = members / population,
         lds_per_capita = lds_prop * 100000)

data("WaffleDivorce")
lds_divorce <- WaffleDivorce %>%
  as_tibble() %>%
  select(Location, Divorce, Marriage, MedianAgeMarriage) %>%
  left_join(select(lds, state, lds_per_capita),
            by = c("Location" = "state")) %>%
  mutate(lds_per_capita = log(lds_per_capita)) %>%
  mutate(across(where(is.numeric), standardize)) %>%
  filter(!is.na(lds_per_capita))

lds_divorce
```

```
# A tibble: 49 x 5
   Location    Divorce Marriage MedianAgeMarriage lds_per_capita
   <chr>         <dbl>    <dbl>              <dbl>          <dbl>
 1 Alabama       1.65    0.0226            -0.606         -0.423
 2 Alaska        1.54    1.55              -0.687          1.21
 3 Arizona       0.611   0.0490            -0.204          1.42
 4 Arkansas      2.09    1.66              -1.41          -0.123
 5 California   -0.927  -0.267              0.600          0.409
 6 Colorado      1.05    0.892             -0.285          0.671
 7 Connecticut  -1.64   -0.794              1.24          -0.909
 8 Delaware     -0.433   0.786              0.439         -0.693
 9 Florida      -0.652  -0.820              0.278         -0.466
10 Georgia       0.995   0.523             -0.124         -0.375
# i 39 more rows
```

Let's create the model

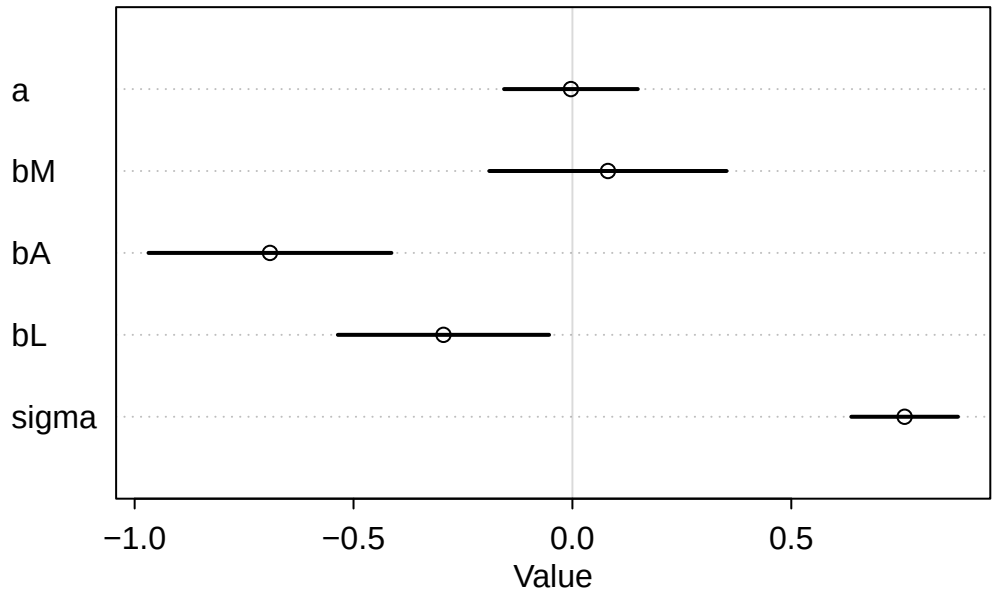
```r
m <- quap(
  alist(
    Divorce ~ dnorm(mu, sigma),
    mu <- a + bM * Marriage + bA * MedianAgeMarriage + bL * lds_per_capita,
    a ~ dnorm(0,0.2),
    c(bM, bA, bL) ~ dnorm(0,0.5),
    sigma ~ dexp(1)
```

```
  ), data=lds_divorce)

print(precis(m))
```

```
              mean         sd       5.5%       94.5%
a     -0.003415348 0.09558462 -0.1561780  0.14934734
bM     0.081157181 0.16955160 -0.1898190  0.35213339
bA    -0.690926343 0.17360606 -0.9683824 -0.41347033
bL    -0.294559603 0.15094072 -0.5357920 -0.05332718
sigma  0.758827075 0.07621448  0.6370216  0.88063253
```

```
plot(precis(m))
```



Exercise 5H1

Let's use `dagitty`

```
library(dagitty)

mad <- dagitty("dag{M->A->D}")
impliedConditionalIndependencies(mad)
```

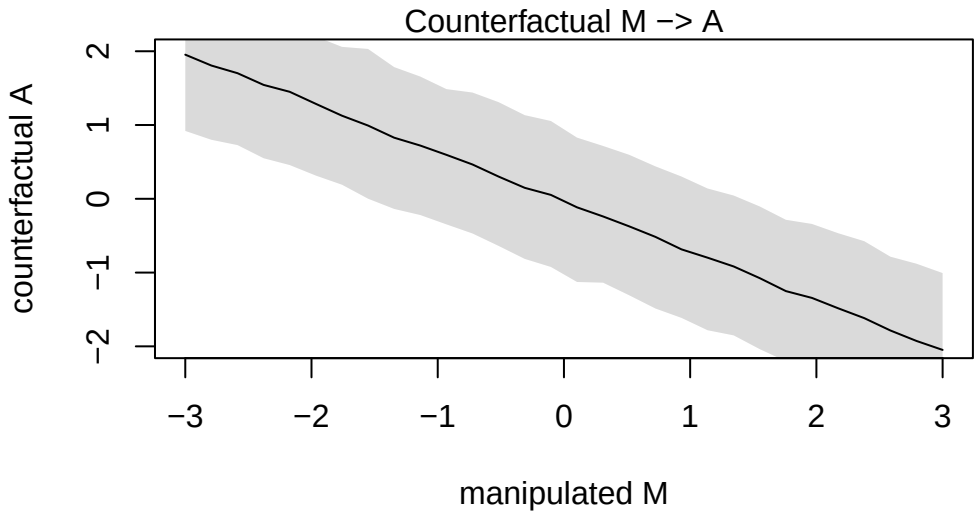
D _||_ M | A

Exercise 5H2

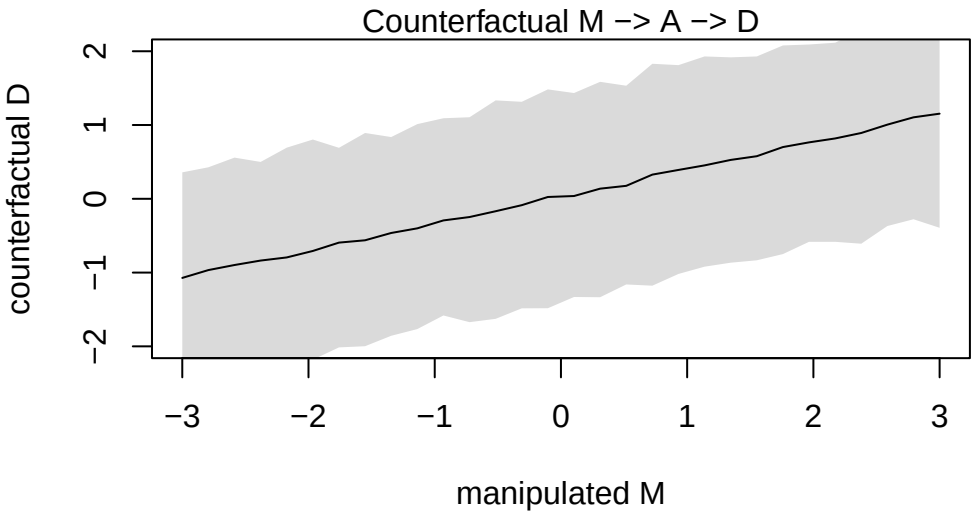
```
m5H2 <- quap(
  alist(
    # A -> D
    Divorce ~ dnorm( muD , sigmaD ),
    muD <- aD + bAD*MedianAgeMarriage,
    # M -> A
    MedianAgeMarriage ~ dnorm( muA , sigmaA ),
    muA <- aA + bMA*Marriage,
    # priors
    c(aD,aA) ~ dnorm(0,0.2),
    c(bAD,bMA) ~ dnorm(0,0.5),
    c(sigmaD,sigmaA) ~ dexp(1)
  ) , data=lds_divorce )
precis(m5H2)
```

```
               mean         sd       5.5%       94.5%
aD       0.003636875 0.09901106 -0.1546019  0.16187568
aA      -0.043444293 0.07765531 -0.1675525  0.08066389
bAD     -0.554240626 0.12184617 -0.7489743 -0.35950691
bMA     -0.669026562 0.08349898 -0.8024741 -0.53557906
sigmaD   0.796052924 0.07956560  0.6688917  0.92321412
sigmaA   0.589724621 0.05910697  0.4952603  0.68418897
```

```
M_seq <- seq( from=-3 , to=3 , length.out=30 )
sim_dat <- data.frame( Marriage=M_seq )
s <- sim( m5H2 , data=sim_dat , vars=c("MedianAgeMarriage","Divorce") )
plot( sim_dat$Marriage , colMeans(s$MedianAgeMarriage) , ylim=c(-2,2) , type="l" ,
xlab="manipulated M" , ylab="counterfactual A" )
shade( apply(s$MedianAgeMarriage,2,PI) , sim_dat$Marriage )
mtext( "Counterfactual M -> A" )
```




```
plot( sim_dat$Marriage , colMeans(s$D) , ylim=c(-2,2) , type="l" ,
xlab="manipulated M" , ylab="counterfactual D" )
shade( apply(s$Divorce,2,PI) , sim_dat$Marriage )
mtext( "Counterfactual M -> A -> D" )
```


Counterfactual M −> A −> D
counterfactual D
−2
−1
0
1
2
−3
−2
−1
0
1
2
3
manipulated M