SR HW 5

Emilio Horner

2022-10-05

```
library(rethinking)
```

```
## Loading required package: rstan
```

```
## Loading required package: StanHeaders
```

```
## Loading required package: ggplot2
```

```
## rstan (Version 2.21.5, GitRev: 2e1f913d3ca3)
```

```
## For execution on a local, multicore CPU with excess RAM we recommend calling
## options(mc.cores = parallel::detectCores()).
## To avoid recompilation of unchanged Stan programs, we recommend calling
## rstan_options(auto_write = TRUE)
```

```
## Loading required package: cmdstanr
```

```
## This is cmdstanr version 0.5.3
```

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

```
## - Use set_cmdstan_path() to set the path to CmdStan
```

```
## - Use install_cmdstan() to install CmdStan
```

```
## Loading required package: parallel
```

```
## rethinking (Version 2.23)
```

```
##
## Attaching package: 'rethinking'
```

```
## The following object is masked from 'package:rstan':
##
##     stan
```

```
## The following object is masked from 'package:stats':
##
##     rstudent
```

5e1.

Options 2, 3, and 4 are all examples of multiple regressions

5e2.

μi = α + βxXi + βzZi

X is animal diversity and Z is plant diversity

5e3.

μi = α + βxXi + βzZi

Here X would be funding and Z would be size of laboratory. Both the slopes would be positive because there's a positive correlation

5e4.

All of them except number 2 are inferentially equivalent

5m1.

```
X <- 100
pizza <- rnorm(n = 100, mean = 0, sd = 6)
dogs <- rnorm(n = X, mean = pizza, sd = 2)
carowner <- rnorm(n = X, mean = pizza, sd = 1)
d <- data.frame(carowner, dogs, pizza)
```

My spurious correlation is that that number of pizzas eaten per year is correlated to dogs owned. But when carownership rate is factored in it is revealed to be a spurious correlation.

```
Y <- map(
  alist(
    carowner ~ dnorm(mu, sigma),
    mu <- a + bo * dogs + bi * pizza,
    a ~ dnorm(0, 5),
    bo ~ dnorm(0, 5),
    bi ~ dnorm(0, 5),
    sigma ~ dunif(0, 5)
  ),
  data = d
)
precis(Y)
```

```
##             mean         sd       5.5%     94.5%
## a     0.22007500 0.09539652 0.06761293 0.3725371
## bo    0.09472879 0.04883564 0.01668001 0.1727776
## bi    0.93162743 0.05210270 0.84835726 1.0148976
## sigma 0.93873678 0.06637845 0.83265120 1.0448224
```

5m2.

```
N <- 500
rho <- 0.4
cars <- rnorm(n = N, mean = 0, sd = 1)
carcrashes <- rnorm(n = N, mean = rho * cars, sd = sqrt(1 - rho^2))
happiness <- rnorm(n = N, mean = cars - carcrashes, sd = 1)
e <- data.frame(happiness, cars, carcrashes)
```

Rate of car ownership makes people happy but getting in car crashes makes people less happy.

```
A <- map(
  alist(
    happiness ~ dnorm(mu, sigma),
    mu <- a + ba * cars + bi * carcrashes,
    a ~ dnorm(0, 5),
    ba ~ dnorm(0, 5),
    bi ~ dnorm(0, 5),
    sigma ~ dunif(0, 5)
  ),
  data = e
)
precis(A)
```

```
##              mean         sd        5.5%      94.5%
## a      0.03084952 0.04343642 -0.03857027  0.1002693
## ba     1.03983125 0.04579518  0.96664172  1.1130208
## bi    -1.01420027 0.04821452 -1.09125638 -0.9371442
## sigma  0.96863867 0.03063093  0.91968452  1.0175928
```

5m3.

If more people are getting divorced there's a higher marriage rate probably people it means people are getting remarried. So people count as getting married more than once. This could be tested by adding the variable "first time marriages".

5m4.

Use LDS per state as a predictor variable, predicting divorce rate using marriage rate, median age at marriage, and percent LDS population

```
library(rethinking)
data("WaffleDivorce")
```

```
mormondata <- WaffleDivorce
mormondata$MormonPerState <- c(0.0077, 0.0453, 0.0610, 0.0104, 0.0194, 0.0270, 0.0044, 0.0057,
0.0041, 0.0075, 0.0082, 0.0520, 0.2623, 0.0045, 0.0067, 0.0090, 0.0130, 0.0079, 0.0064, 0.0082,
0.0072, 0.0040, 0.0045, 0.0059, 0.0073, 0.0116, 0.0480, 0.0130, 0.0065, 0.0037, 0.0333, 0.0041,
0.0084, 0.0149, 0.0053, 0.0122, 0.0372, 0.0040, 0.0039, 0.0081, 0.0122, 0.0076, 0.0125, 0.6739,
0.0074, 0.0113, 0.0390, 0.0093, 0.0046, 0.1161)
```

```
W <- map(
  alist(
    Divorce ~ dnorm(mu, sigma),
    mu <- a + bm * Marriage + ba * MedianAgeMarriage + bM * MormonPerState,
    a ~ dnorm(10, 20),
    bm ~ dnorm(0, 10),
    ba ~ dnorm(0, 10),
    bM ~ dnorm(0, 10),
    sigma ~ dunif(0, 5)
  ),
  data = mormondata
)
precis(W)
```

```
##              mean         sd        5.5%      94.5%
## a     35.50202100 6.59927705 24.95510168 46.0489403
## bm     0.02481754 0.07355366 -0.09273541  0.1423705
## ba    -1.00240671 0.21434922 -1.34497817 -0.6598353
## bM    -5.82155958 2.19713140 -9.33299992 -2.3101192
## sigma  1.34248453 0.13480372  1.12704215  1.5579269
```

I get that there’s a negative corelation of -5.84 for the mean of divorce rates in places with more Mormons. I think this means that the Mormons don’t get divorced as much.

5m5.

An example of how buying a notebook impacts your grades. One explanation could be better daily scheduling organization. Anter could be that the act of taking notes helps form memories.

You could test this by adding another variable in addition to grades (outcome variable) and notebook ownership (predictor variable) such as percent of classes in which someone wrote down notes (predictor variable)