Assignment 2 - Methods 4

Laurits Lyngbaek + study group 5

2025-03-18

Note:

This assignment has been completed in the study group. The members are: Andreas Møldrup Holst, Zofia Radwanska, Anne Keena, Maja Gade Mortensen and Eva Quiring. The responsible person is mentioned in the heading and is responsible until the next name is mentioned.

Second assignment

The second assignment uses chapter 3, 5 and 6. The focus of the assignment is getting an understanding of causality.

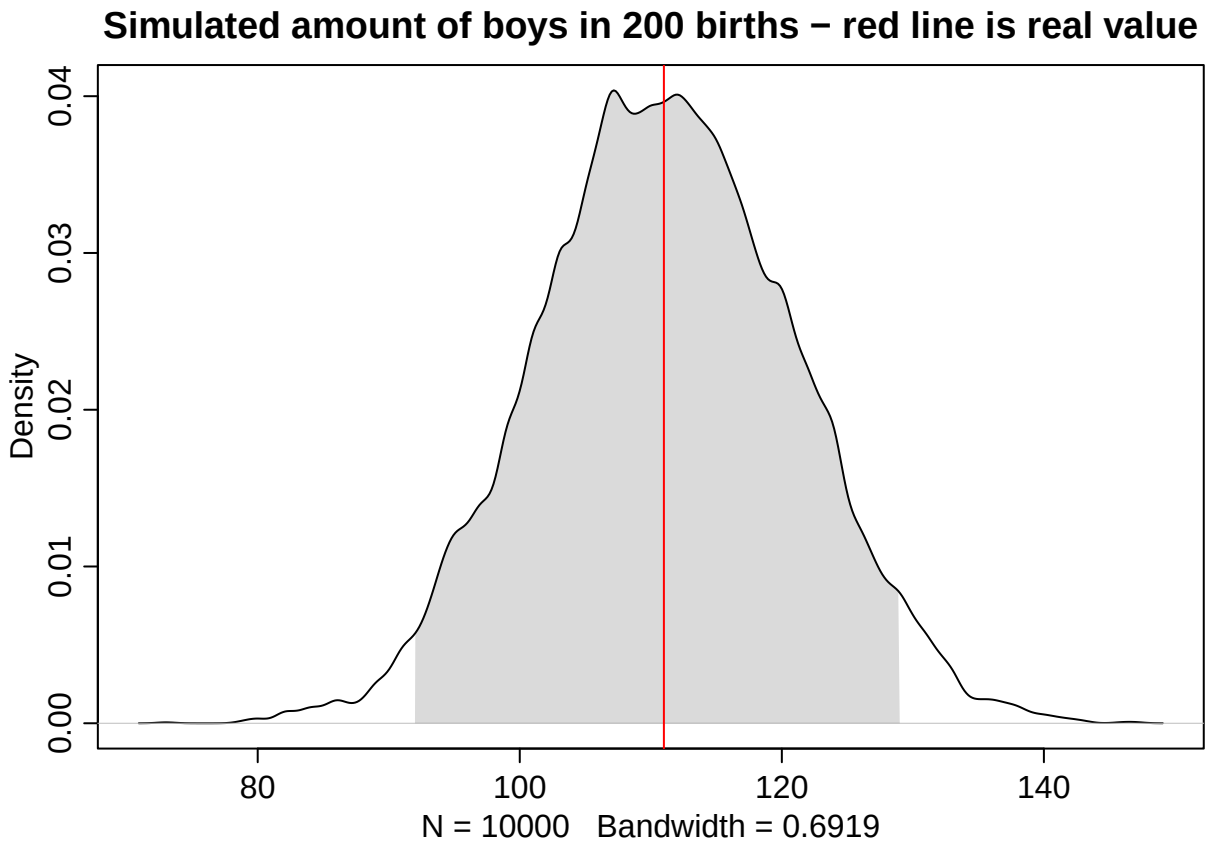
Chapter 3: Causal Confussion - Maja

Reminder: We are tying to estimate the probability of giving birth to a boy I have pasted a working solution to questions 6.1-6.3 so you can continue from here:)

3H3 Use rbinom to simulate 10,000 replicates of 200 births. You should end up with 10,000 numbers, each one a count of boys out of 200 births. Compare the distribution of predicted numbers of boys to the actual count in the data (111 boys out of 200 births).

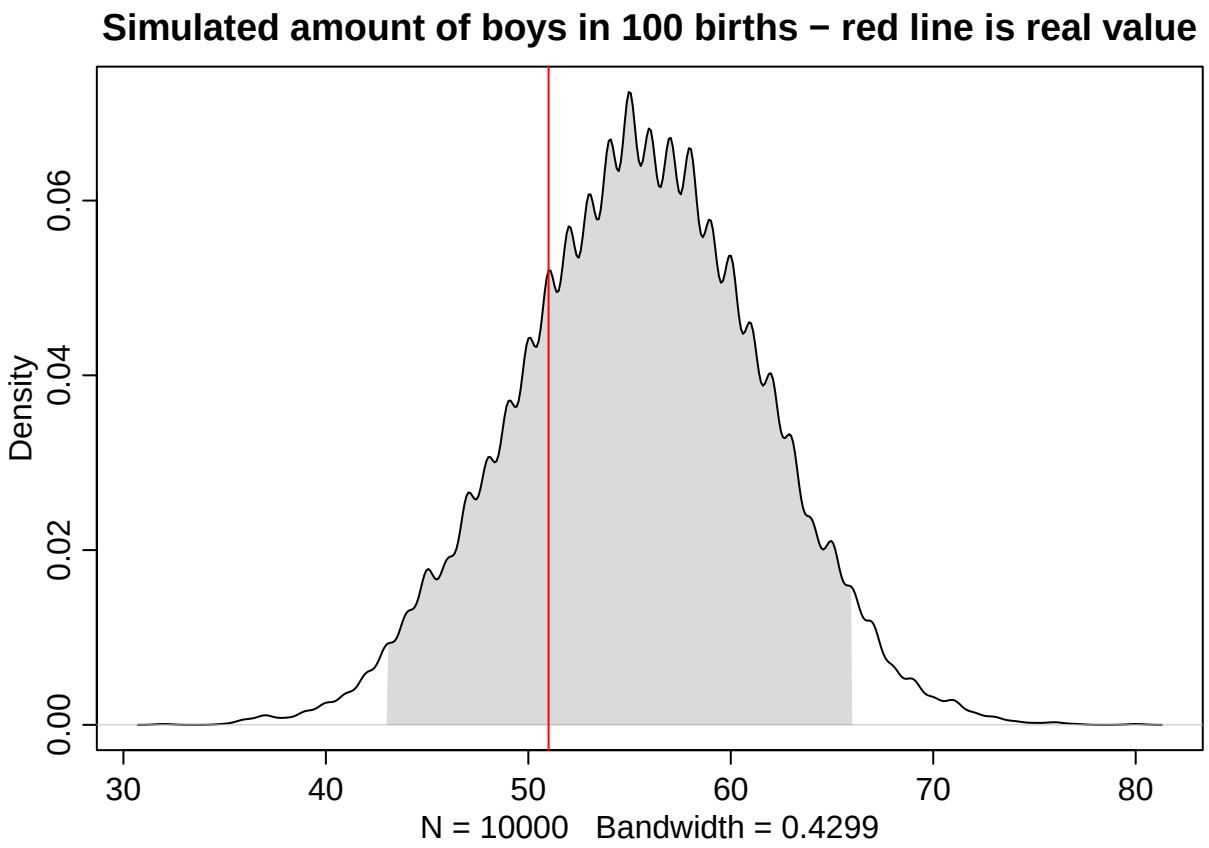
```
# 3H1
# Find the posterior probability of giving birth to a boy:
pacman::p_load(rethinking)
data(homeworkch3)
set.seed(1)
W <- sum(birth1) + sum(birth2)
N <- length(birth1) + length(birth2)
p_grid <-seq(from =0, to = 1, len =1000)
prob_p <- rep(1,1000)
prob_data <- dbinom(W,N,prob=p_grid)
posterior <-prob_data * prob_p
posterior <- posterior / sum(posterior)

# 3H2
# Sample probabilities from posterior distribution:
samples <- sample (p_grid, prob = posterior, size =1e4, replace =TRUE)


# 3H3
# Simulate births using sampled probabilities as simulation input, and check if they allign with real v
```

```r
simulated_births <- rbinom(n = 1e4, size = N, prob = samples)
rethinking::dens(simulated_births,show.HPDI = 0.95)
abline(v=W, col="red")
title("Simulated amount of boys in 200 births - red line is real value")
```



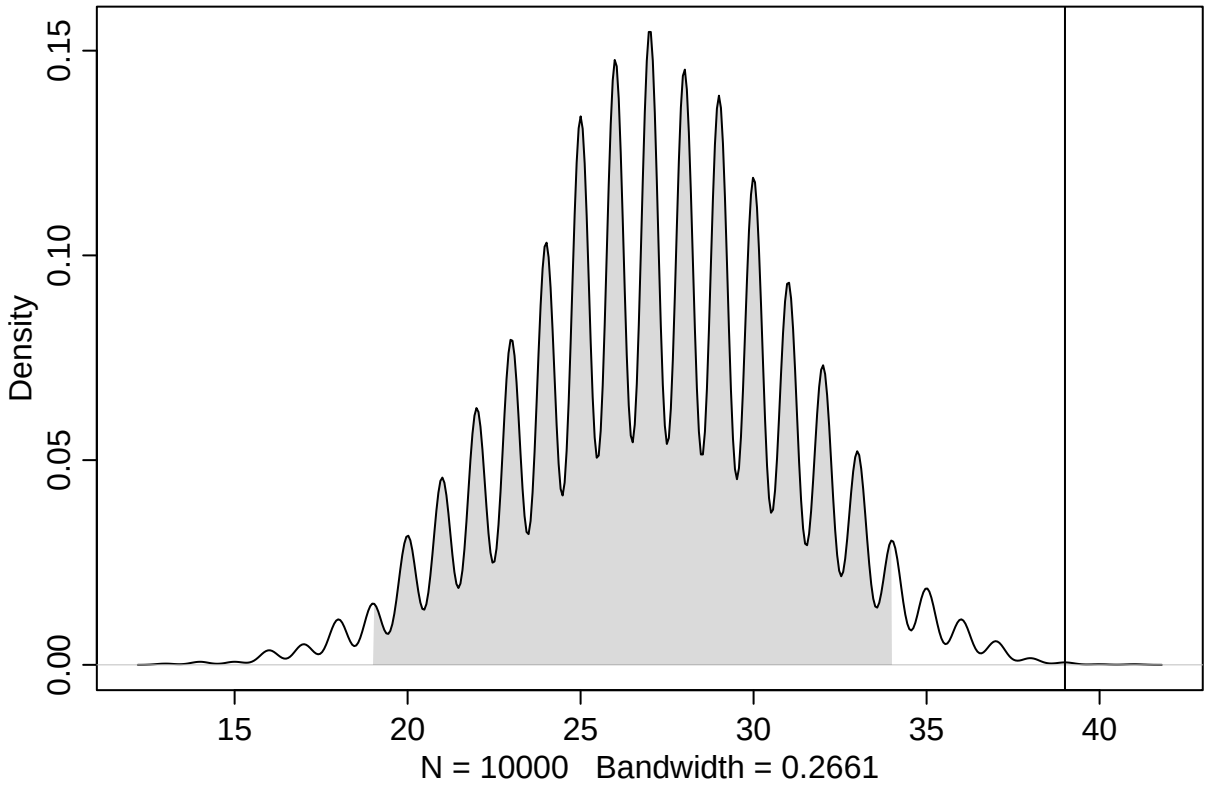
3H4. Now compare 10,000 counts of boys from 100 simulated first borns only to the number of boys in the first births, birth1. How does the model look in this light?

```r
set.seed(1)
W4 <- sum(birth1)
N4 <- length(birth1)


# Simulate births using sampled probabilities as simulation input, and check if they allign with real v
simulated_births4 <- rbinom(n = 1e4, size = N4, prob = samples)
rethinking::dens(simulated_births4,show.HPDI = 0.95)
abline(v=W4, col="red")
title("Simulated amount of boys in 100 births - red line is real value")
```



3H5. The model assumes that sex of first and second births are independent. To check this assumption, focus now on second births that followed female first borns. Compare 10,000 simulated counts of boys to only those second births that followed girls. To do this correctly, you need to cound the number of first borns who were girls and simulate that many births, 10,000 times. Compare the counts of boys in your simulations to the actual observed count of boys following girls. How does the model look in this light? Any guesses what is going on in these data?

```
# Number of first born girls
counter <- 0
for (i in birth1){
  if (i == 0){
    counter <- counter + 1
  }
}
#counter #49

#The births where girls are born first
girls_first <- birth2[birth1 == 0]

boys_after_girls <- sum(girls_first) #39 boys

#Simulate boys folowing girls
  sim_boys_girls <- rbinom(10000, size = counter, prob = samples)
rethinking::dens(sim_boys_girls,show.HPDI = 0.95)
abline(v = boys_after_girls)
```



Answer: The model is not doing very well since it doesn't reflect the real scenario (the vertical line is not even close to the peak of the distribution). This means that there might be a causal relationship between the gender of the second birth and the first birth.

Chapter 5: Spurrious Correlations - Maja

Start of by checking out all the spurrious correlations that exists in the world. Some of these can be seen on this wonderfull website: https://www.tylervigen.com/spurious/random All the medium questions are only asking you to explain a solution with words, but feel free to simulate the data and prove the concepts.

5M1. Invent your own example of a spurious correlation. An outcome variable should be correlated with both predictor variables. But when both predictors are entered in the same model, the correlation between the outcome and one of the predictors should mostly vanish (or at least be greatly reduced).

Answer: Outcome: It's good weather Preditcor 1: There's sun Predictor 2: People are eating ice cream

One would expect that the second predictor mostly vanish when both predictors are entered into a model.

5M2. Invent your own example of a masked relationship. An outcome variable should be correlated with both predictor variables, but in opposite directions. And the two predictor variables should be correlated with one another.

Answer: Outcome: Ratings of a restaurant Predictor 1: Food left at the end of the day Predictor 2: Number of people eating at the restaurant at the particular day

One would expect that the ratings of the restaurant correlate with both the food that is left (if more food is left, we expect that the restaurant get worse ratings) in addition, we assume that the more people who visit the restaurant the better the ratings.

5M3. It is sometimes observed that the best predictor of fire risk is the presence of firefighters— States and localities with many firefighters also have more fires. Presumably firefighters do not cause fires. Nevertheless, this is not a spurious correlation. Instead fires cause firefighters. Consider the same reversal of causal inference in the context of the divorce and marriage data. How might a high divorce rate cause a higher marriage rate? Can you think of a way to evaluate this relationship, using multiple regression

Answer: In order to have high divorce rate many people need to get married (a certain percentage get divorced). In a classical multiple regression we set one of them as a predictor and the other one as the outcome variable: marriage rate ~ a + divorce rate * b However, this doesn't show if divorce depends on marriage or vice versa. If we include other predictors such as age when married, gender, or if it's a first time marriage, might help understanding the relationship between divorce rate and marriage rate.

5M5. One way to reason through multiple causation hypotheses is to imagine detailed mechanisms through which predictor variables may influence outcomes. For example, it is sometimes argued that the price of gasoline (predictor variable) is positively associated with lower obesity rates (outcome variable). However, there are at least two important mechanisms by which the price of gas could reduce obesity. First, it could lead to less driving and therefore more exercise. Second, it could lead to less driving, which leads to less eating out, which leads to less consumption of huge restaurant meals. Can you outline one or more multiple regressions that address these two mechanisms? Assume you can have any predictor data you need.

Answer: obesity ~ a + gasoline_price x b1 + number_of_people_exercising x b2 + time_people_exercise x b3 + consumption_of_huge_restaurant_meals x b4

Chapter 5: Foxes and Pack Sizes - Anne, Andreas, Eva

All five exercises below use the same data, data(foxes) (part of rethinking).84 The urban fox (Vulpes vulpes) is a successful exploiter of human habitat. Since urban foxes move in packs and defend territories, data on habitat quality and population density is also included. The data frame has five columns: (1) group: Number of the social group the individual fox belongs to (2) avgfood: The average amount of food available in the territory (3) groupsize: The number of foxes in the social group (4) area: Size of the territory (5) weight: Body weight of the individual fox

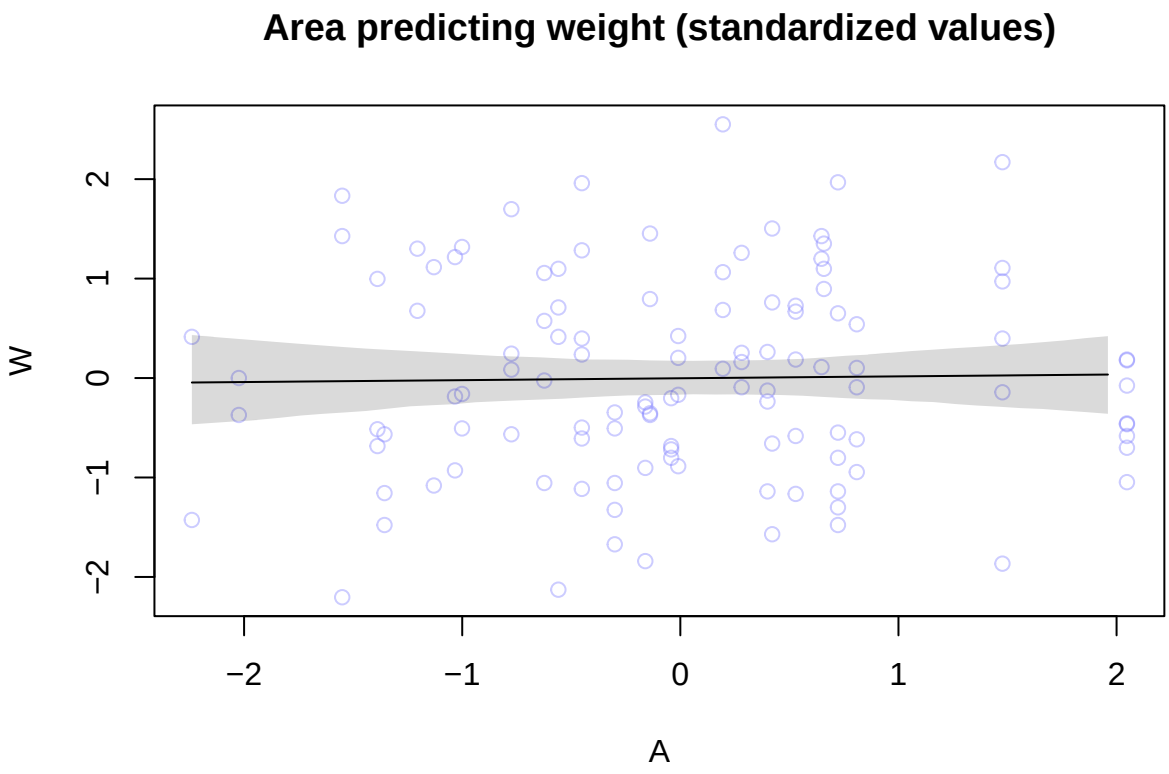
Andreas

5H1. Fit two bivariate Gaussian regressions, using quap: (1) body weight as a linear function of territory size (area), and (2) body weight as a linear function of groupsize. Plot the results of these regressions, displaying the MAP regression line and the 95% interval of the mean. Is either variable important for predicting fox body weight?

```
library(rethinking)
pacman::p_load(tidybayes, tibble)
data(foxes)
fox_data <- foxes
fox_data$A <- standardize(fox_data$area)
fox_data$W <- standardize(fox_data$weight)
fox_data$AF <- standardize(fox_data$avgfood)
fox_data$G <- standardize(fox_data$group)
fox_data$GS <- standardize(fox_data$groupsize)
```

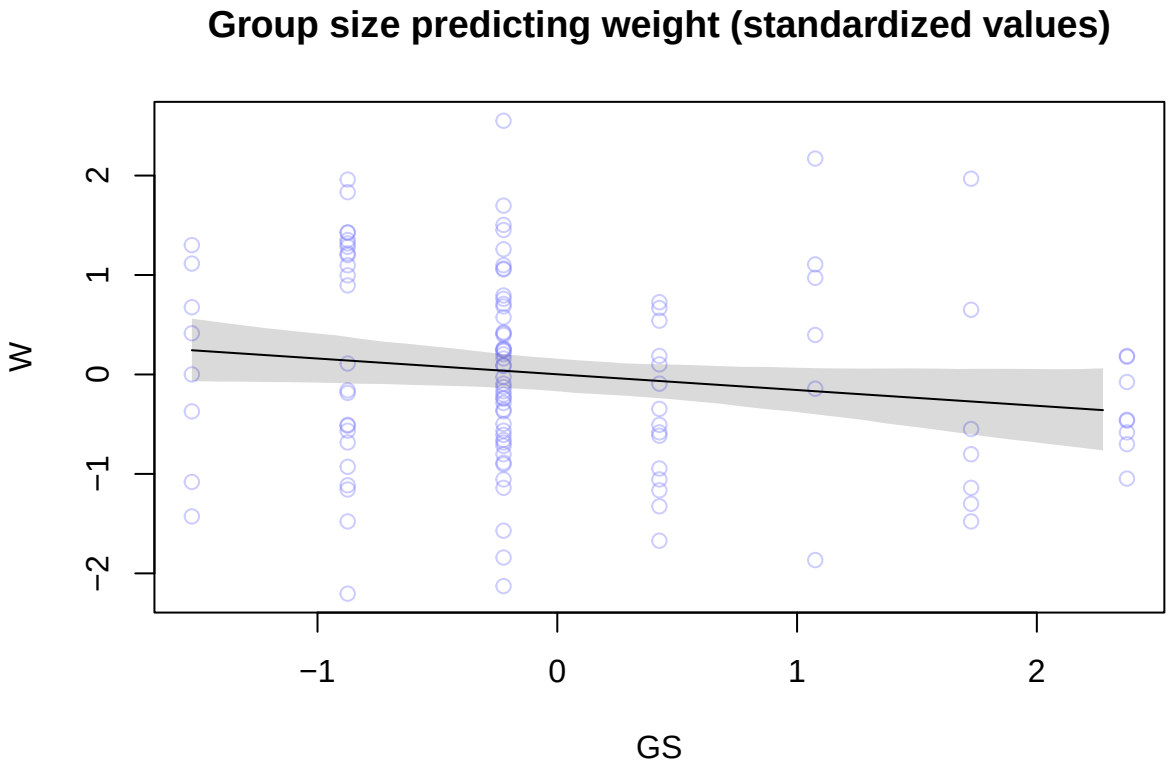
```r
# Create model 1

model_1 <- quap(alist(
    W ~ dnorm(mu, sigma),
    mu <- a + bA * A,
    a ~ dnorm(0, 0.2),
    bA ~ dnorm(0, 0.5),
    sigma ~ dexp(1)
  ), data = fox_data)

#precis(model_1)
```

```r
#Plot
plot(W ~ A , data=fox_data, col=col.alpha(rangi2,0.4),
     main = "Area predicting weight (standardized values)")

#Add line
# Making an area sequency to make predictions
area_seq <- seq(from = min(fox_data$A), to = max(fox_data$A), by = 0.1)
mu_1 <- link(model_1, data = data.frame(A = area_seq))
mu_mean_1 <- apply(mu_1, 2, mean)
lines(area_seq, mu_mean_1)


#95% interval
mu_PI <- apply(mu_1, 2, PI, prob=0.95)
shade(mu_PI, area_seq)
```




```
# Create model 2
model_2 <- quap(alist(
    W ~ dnorm(mu, sigma),
    mu <- a + bGS * GS,
    a ~ dnorm(0, 0.2),
    bGS ~ dnorm(0, 0.5),
    sigma ~ dexp(1)
  ), data = fox_data)

#precis(model_2)
```

```
#Plot - SOMETHING IS WRONG
plot(W ~ GS , data=fox_data, col=col.alpha(rangi2,0.4),
     main = "Group size predicting weight (standardized values)")

#Add line
# Making an area sequency to make predictions
GS_seq <- seq(from = min(fox_data$GS), to = max(fox_data$GS), by = 0.1)
mu_2 <- link(model_2, data = data.frame(GS = GS_seq))
mu_mean_2 <- apply(mu_2, 2, mean)
lines(GS_seq, mu_mean_2)


#95% interval
mu_PI_2 <- apply(mu_2, 2, PI, prob=0.95)
shade(mu_PI_2, GS_seq)
```



Answer: The size of the area doesn't seem to influence the weight of the foxes. In addition, there might be a little tendency that group size can predict the weight of the foxes (the smaller the group the heavier the foxes).

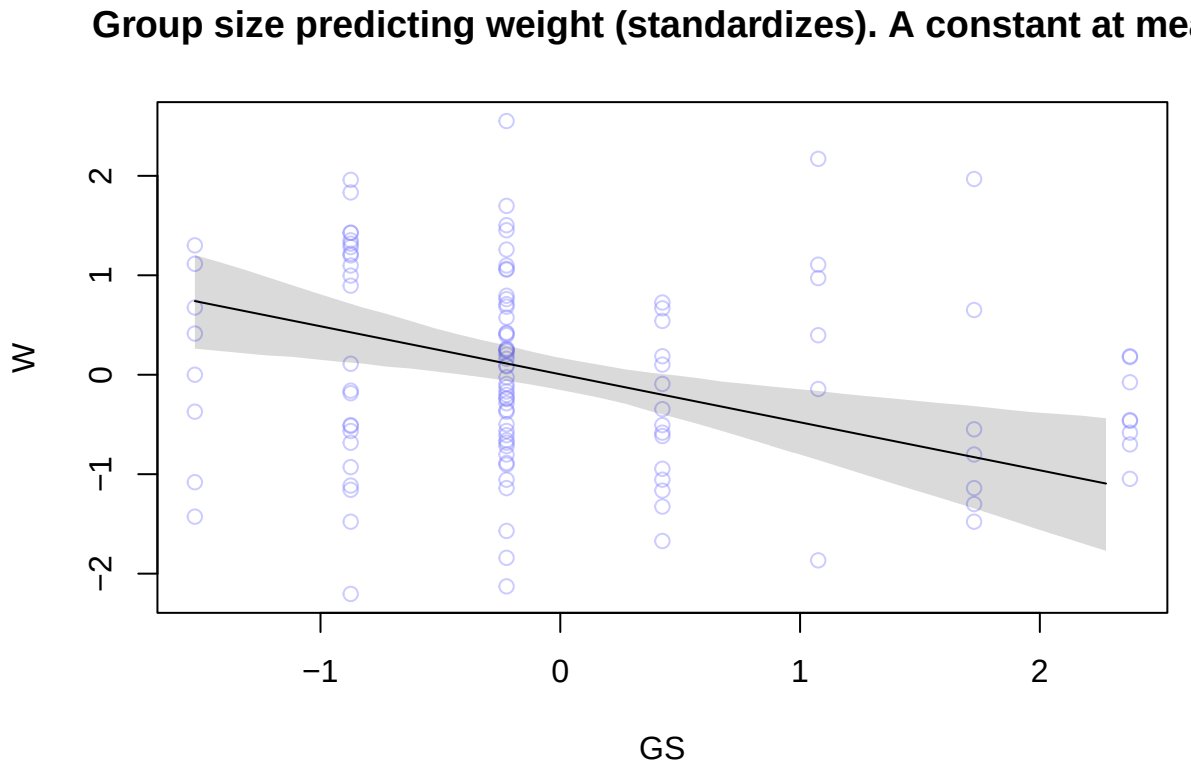
Eva

5H2. Now fit a multiple linear regression with weight as the outcome and both area and groupsize as predictor variables. Plot the predictions of the model for each predictor, holding the other predictor constant at its mean. What does this model say about the importance of each variable? Why do you get different results than you got in the exercise just above?

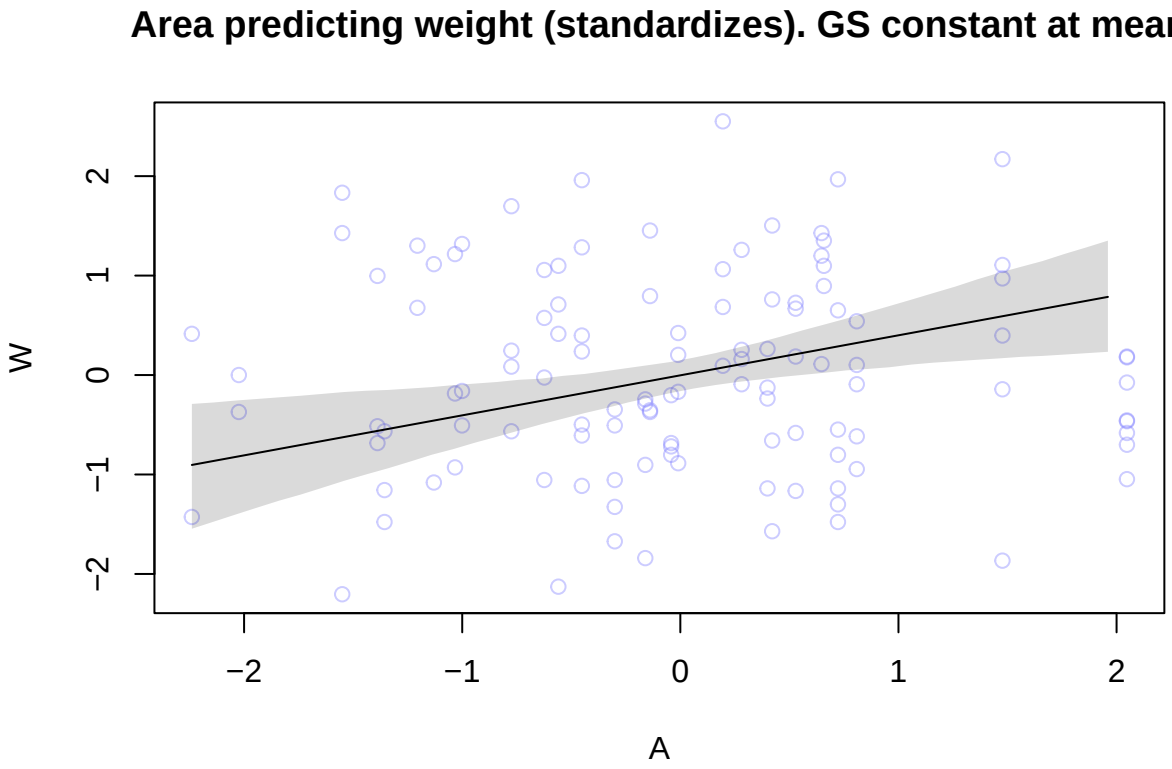
```
model_3 <- quap(alist(
    W ~ dnorm(mu, sigma),
    mu <- a + bGS * GS + bA * A,
    a ~ dnorm(0, 0.2),
    bGS ~ dnorm(0, 0.5),
    bA ~ dnorm(0, 0.5),
    sigma ~ dexp(1)
  ), data = fox_data)
```

```
mean_A <- mean(fox_data$A)
```

```r
plot(W ~ GS , data=fox_data, col=col.alpha(rangi2,0.4),
     main = "Group size predicting weight (standardizes). A constant at mean")

mu_GS <- link(model_3, data = data.frame(GS = GS_seq, A = mean_A))
mu_mean_GS <- apply(mu_GS, 2, mean)
lines(GS_seq, mu_mean_GS)

#95% interval
mu_PI_GS <- apply(mu_GS, 2, PI, prob=0.95)
shade(mu_PI_GS, GS_seq)
```




```r
mean_GS <- mean(fox_data$GS)

plot(W ~ A , data=fox_data, col=col.alpha(rangi2,0.4),
     main = "Area predicting weight (standardizes). GS constant at mean")

mu_A <- link(model_3, data = data.frame(A = area_seq, GS = mean_GS))
mu_mean_A <- apply(mu_A, 2, mean)
lines(area_seq, mu_mean_A)

#95% interval
mu_PI_A <- apply(mu_A, 2, PI, prob=0.95)
shade(mu_PI_A, area_seq)
```



Answer: In this model it seems like both of the variables are better at predicting weight (steeper slope). The predictors probably covary, but by keeping them constant at the mean we can observe how they interfere with the outcome variable weight of the foxes.

Anne

5H3. Finally, consider the avgfood variable. Fit two more multiple regressions: (1) body weight as an additive function of avgfood and groupsize, and (2) body weight as an additive function of all three variables, avgfood and groupsize and area. Compare the results of these models to the previous models you've fit, in the first two exercises. (a) Is avgfood or area a better predictor of body weight? If you had to choose one or the other to include in a model, which would it be? Support your assessment with any tables or plots you choose. (b) When both avgfood or area are in the same model, their effects are reduced (closer to zero) and their standard errors are larger than when they are included in separate models. Can you explain this result?

```
# Model 4:
model_4 <- quap(alist(
    W ~ dnorm(mu, sigma),
    mu <- a + Bf*AF + Bg*GS,
    a ~ dnorm(0, 0.2),
    Bf ~ dnorm(0, 0.5),
    Bg ~ dnorm(0, 0.5),
    sigma ~ dexp(1)
  ), data = fox_data)
```

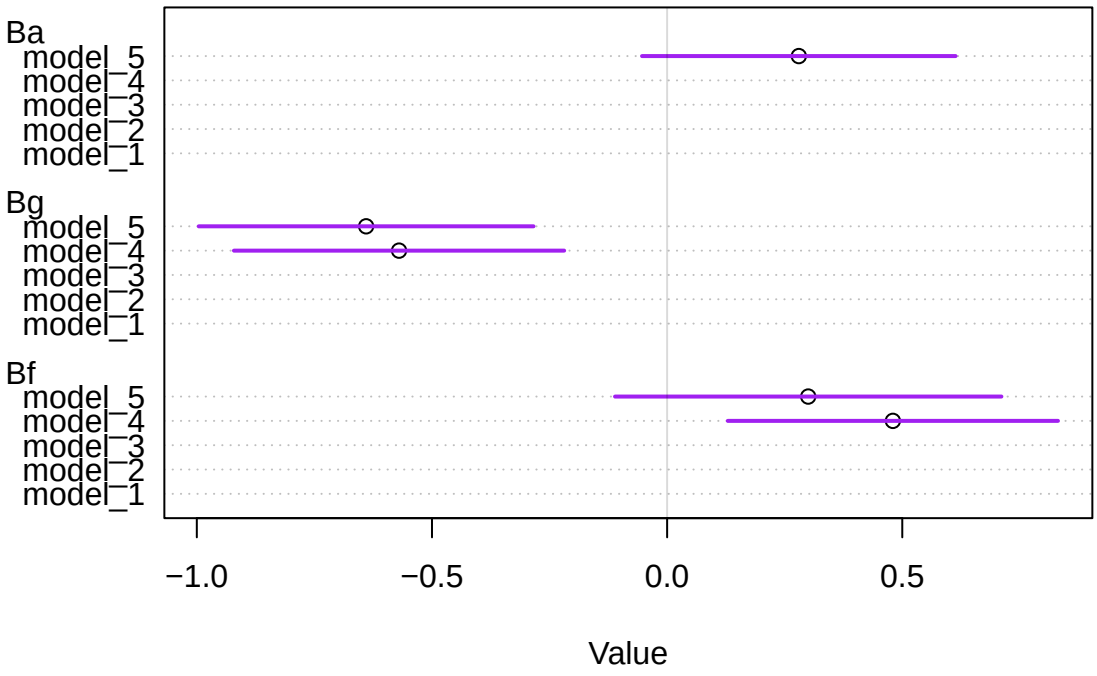
```
# Model 5:
model_5 <- quap(alist(
    W ~ dnorm(mu, sigma),
    mu <- a + Bf*AF + Bg*GS + Ba*A,
    a ~ dnorm(0, 0.2),
    Bf ~ dnorm(0, 0.5),
    Bg ~ dnorm(0, 0.5),
    Ba ~ dnorm(0, 0.5),
    sigma ~ dexp(1)
  ), data = fox_data)

precis(model_4)
```

```
##                mean         sd       5.5%      94.5%
## a     -1.892222e-05 0.08013651 -0.1280925  0.1280547
## Bf     4.772744e-01 0.17911980  0.1910064  0.7635424
## Bg    -5.735452e-01 0.17913831 -0.8598428 -0.2872476
## sigma  9.420221e-01 0.06174898  0.8433353  1.0407089
```

```
precis(model_5)
```

```
##                mean         sd         5.5%      94.5%
## a      5.613667e-06 0.07935947 -0.126826140  0.1268374
## Bf     2.970100e-01 0.20959370 -0.037961164  0.6319813
## Bg    -6.398319e-01 0.18160739 -0.930075592 -0.3495882
## Ba     2.783746e-01 0.17010658  0.006511385  0.5502377
## sigma  9.311709e-01 0.06099424  0.833690343  1.0286515
```

```
plot(coeftab(model_1, model_2 , model_3, model_4, model_5),
     pars=c("Ba","Bg", "Bf"),
     col = c("purple"))
```



Answer (a): In model_4, avgfood shows a strong positive relationship with weight (Bf = 0.48). groupsize also contributes to explaining variation in weight. When area is added as a predictor in model_5, the avgfood coefficient drops to 0.3 and groupsize changes from from -0.57 to -0.65). Adding area as an additional predictor shows a positive effect (Ba = 0.28). When comparing these results to previous models, we see that both avgfood and area are related to body weight but avgfood appears to be a stronger predictor and shows a clearer relationship with body weight.

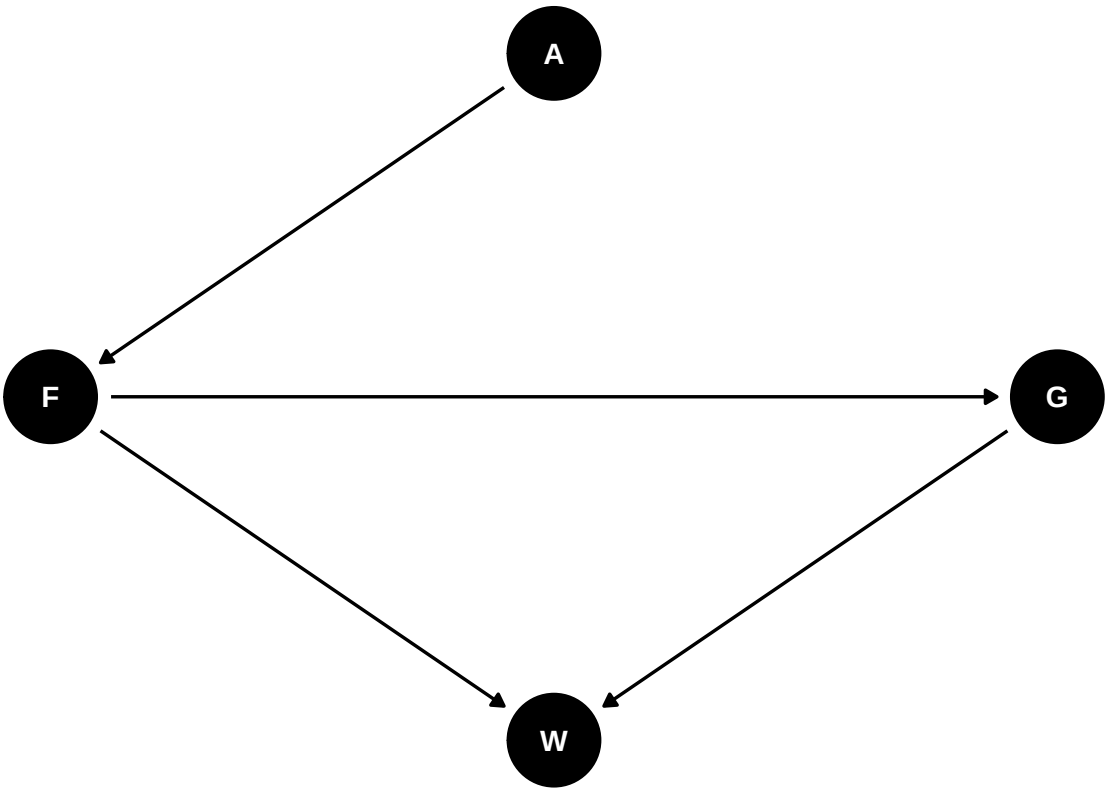
Answer (b): When both avgfood and area are included in the same model, their coefficients move closer to zero and their standard errors increase because of multicollinearity, where the two predictors are strongly correlated. The model then has difficulty distinguishing the specific contribution of each predictor since they share a lot of the same information about body weight. Each variable's estimated effect then becomes smaller and we see more uncertainty around the estimates. This tells us that it would be better only to include one of the predictors in the final model.

Defining our theory with explicit DAGs Assume this DAG as an causal explanation of fox weight:

```
#install.packages("igraph")
pacman::p_load(dagitty,
               ggdag)
dag <- dagitty('dag {
A[pos="1.000,0.500"]
F[pos="0.000,0.000"]
G[pos="2.000,0.000"]
W[pos="1.000,-0.500"]
A -> F
```

```
F -> G
F -> W
G -> W
}')
```

```
# Plot the DAG
ggdag::ggdag(dag, layout = "circle")+
  theme_dag()
```



where A is area, F is avgfood,G is groupsize, and W is weight.

Using what you know about DAGs from chapter 5 and 6, solve the following three questions:

1) Estimate the total causal influence of A on F. What effect would increasing the area of a territory have on the amount of food inside of it?

```
library(rethinking)
pacman::p_load(tidybayes, tibble)
data(foxes)
fox_data <- foxes
fox_data$A <- standardize(fox_data$area)
fox_data$W <- standardize(fox_data$weight)
fox_data$AF <- standardize(fox_data$avgfood)
fox_data$G <- standardize(fox_data$groupsize)
```

```
# Create model 1

model_1 <- quap(alist(
    AF ~ dnorm(mu, sigma),
    mu <- a + bA * A,
    a ~ dnorm(0, 0.2),
    bA ~ dnorm(0, 0.5),
    sigma ~ dexp(1)
  ), data = fox_data)

precis(model_1)
```

```
##               mean         sd        5.5%      94.5%
## a     1.077778e-07 0.04231377 -0.06762547 0.06762568
## bA    8.764731e-01 0.04332671  0.80722865 0.94571756
## sigma 4.662886e-01 0.03052953  0.41749648 0.51508066
```

Answer: The prior of the slope was at mean = 0 and a SD = 0.5, and now the estimated mean is 0.88 and the SD = 0.04. This means that it's a casual influence. When increasing the area of territory we also increase the amount of average food in it.

Andreas

2) Infer the **total** causal effect of adding food F to a territory on the weight W of foxes. Can you calculate the causal effect by simulating an intervention on food?

```
model_2 <- quap(alist(
    W ~ dnorm(mu, sigma),
    mu <- a + bA * A + bAF * AF,
    a ~ dnorm(0, 0.2),
    bA ~ dnorm(0, 0.5),
    bAF ~ dnorm(0, 0.5),
    sigma ~ dexp(1)
  ), data = fox_data)

precis(model_2)
```

```
##                mean         sd       5.5%     94.5%
## a     -2.027874e-08 0.08334405 -0.1331999 0.1331999
## bA     1.461374e-01 0.17418830 -0.1322491 0.4245240
## bAF   -1.490384e-01 0.17418847 -0.4274252 0.1293484
## sigma  9.874681e-01 0.06444172  0.8844778 1.0904584
```

```
model_no_food <- quap(alist(
    W ~ dnorm(mu, sigma),
    mu <- a + bA * A,
    a ~ dnorm(0, 0.2),
    bA ~ dnorm(0, 0.5),
    sigma ~ dexp(1)
  ), data = fox_data)
```

```
precis(model_no_food)
```

```
##                mean         sd       5.5%     94.5%
## a     7.638167e-08 0.08360862 -0.1336226 0.1336228
## bA    1.883371e-02 0.09089576 -0.1264353 0.1641027
## sigma 9.912654e-01 0.06466637  0.8879160 1.0946147
```

Answer: We can see, that when stratifying F, there is no causal effect between Area and Weight. This can be seen since the mean of the posterior distribution of bA does not really differ from the prior (the SD of the posterior overlaps with 0 which is our prior). This means that area and weight are conditionally independent when we condition on F.

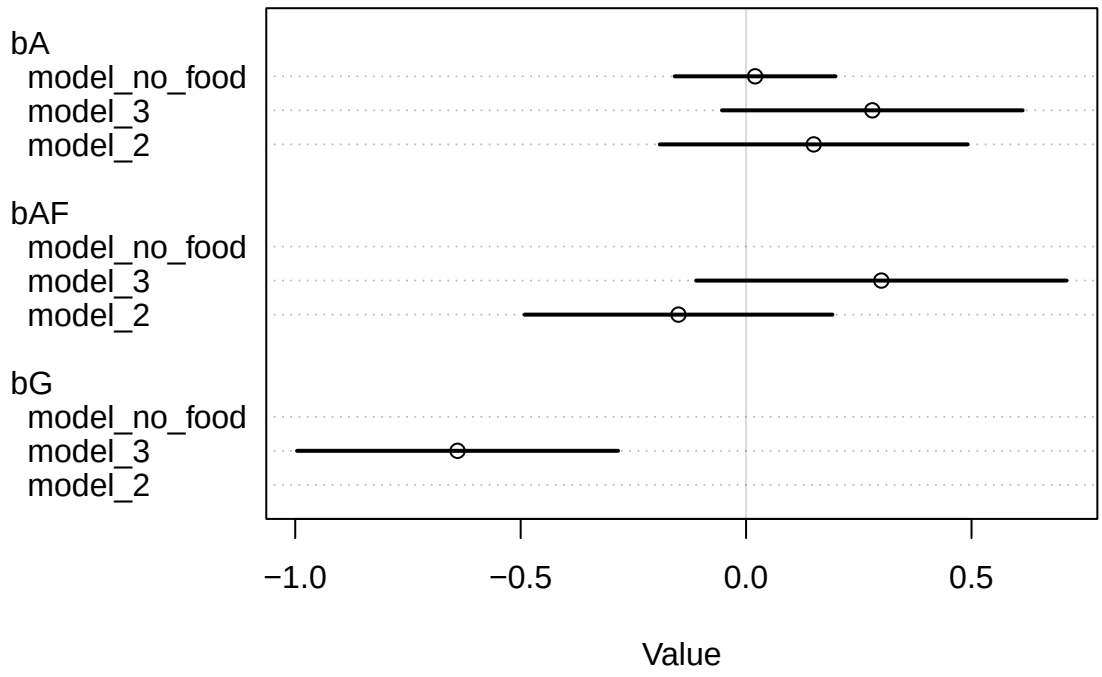
Eva

3) Infer the **direct** causal effect of adding food F to a territory on the weight W of foxes. In light of your estimates from this problem and the previous one, what do you think is going on with these foxes?

```
#Condition on G
model_3 <- quap(alist(
    W ~ dnorm(mu, sigma),
    mu <- a + bA * A + bAF * AF + bG * G,
    a ~ dnorm(0, 0.2),
    bA ~ dnorm(0, 0.5),
    bAF ~ dnorm(0, 0.5),
    bG ~ dnorm(0, 0.5),
    sigma ~ dexp(1)
  ), data = fox_data)

precis(model_3)
```

```
##                mean         sd         5.5%      94.5%
## a     -1.717298e-07 0.07936201 -0.126835988  0.1268356
## bA     2.782378e-01 0.17011227  0.006365513  0.5501100
## bAF    2.968991e-01 0.20960023 -0.038082510  0.6318808
## bG    -6.396194e-01 0.18161483 -0.929874961 -0.3493638
## sigma  9.312063e-01 0.06100008  0.833716403  1.0286962
```

```
plot(coeftab(model_2 , model_3, model_no_food ) , pars=c("bA","bAF", "bG") )
```



Answer: When looking at this model and the DAG, we can see that the weight of the foxes depends on the amount of food. However, this depends on the size of the area. Additionally looking at the DAG, food influences both groups size and weight of the individual foxes. Group size decreases the weight of the foxes - the fewer foxes in a group the heavier they are.

Chapter 6: Investigating the Waffles and Divorces - Zofia

6H1. Use the Waffle House data, data(WaffleDivorce), to find the total causal influence of number of Waffle Houses on divorce rate. Justify your model or models with a causal graph.

```
# Load in the waffleDivorce data
data("WaffleDivorce")
Waffle_data <- WaffleDivorce
Waffle_data$DIV <- standardize(Waffle_data$Divorce)
Waffle_data$SOU <- standardize(Waffle_data$South)
Waffle_data$AGE <- standardize(Waffle_data$MedianAgeMarriage)
Waffle_data$MAR <- standardize(Waffle_data$Marriage)
Waffle_data$WAF <- standardize(Waffle_data$WaffleHouses)
```

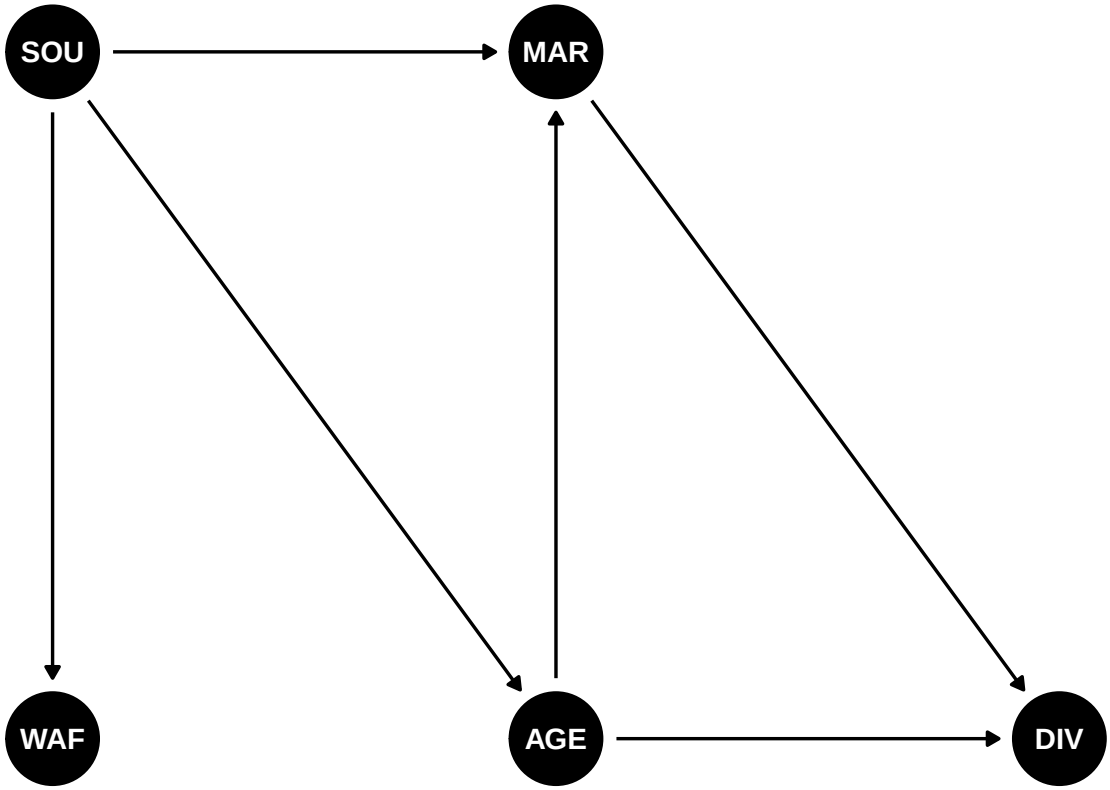
```
model_div <- quap(alist(
    # A -> D <- M
```

```
    DIV ~ dnorm(mu_DIV, sigma_DIV),
    mu_DIV<-aD +bM*MAR +bA*AGE,
    aD ~ dnorm( 0, 0.2),
    bM~ dnorm( 0, 0.5),
    bA~ dnorm( 0, 0.5),
    sigma_DIV~dexp( 1),
    # A -> M <- S
    MAR~dnorm(mu_MAR, sigma_MAR),
    mu_MAR<-aM +bS * SOU +bA*AGE,
    aM~dnorm( 0, 0.2),
    bS~ dnorm( 0, 0.5),
    bA~ dnorm( 0, 0.5),
    sigma_MAR~dexp( 1),
    #S->A
    AGE~dnorm(mu_AGE , sigma_AGE),
    mu_AGE<- aA +bS*SOU,
    aA~ dnorm( 0, 0.2),
    bS~dnorm(0 ,0.5),
    sigma_AGE ~ dexp(1 ),
    # S -> W
    WAF~dnorm(mu_WAF , sigma_WAF),
    mu_WAF<- aW +bSW*SOU,
    aW~ dnorm( 0, 0.2),
    bSW~dnorm(0 ,0.5),
    sigma_WAF ~ dexp(1 )
    ), data = Waffle_data)

precis(model_div)
```

```
##                    mean         sd       5.5%       94.5%
## aD        -3.328444e-06 0.09694851 -0.1549458  0.15493912
## bM        -1.214281e-01 0.12283699 -0.3177453  0.07488917
## bA        -6.951809e-01 0.08174410 -0.8258238 -0.56453805
## sigma_DIV  7.837697e-01 0.07747709  0.6599463  0.90759301
## aM         8.164564e-06 0.08638118 -0.1380456  0.13806197
## bS        -1.364319e-01 0.07858660 -0.2620284 -0.01083531
## sigma_MAR  6.772309e-01 0.06748703  0.5693736  0.78508818
## aA         6.582585e-06 0.11203597 -0.1790485  0.17906171
## sigma_AGE  9.563517e-01 0.09471035  0.8049862  1.10771709
## aW        -1.079465e-05 0.09030043 -0.1443283  0.14430673
## bSW        6.586445e-01 0.10029498  0.4983538  0.81893526
## sigma_WAF  7.156134e-01 0.07090909  0.6022870  0.82893986
```

```
# Create the DAG
dag_6H1 <- dagitty('dag {
    SOU[pos="1, 3"]
    AGE[pos="2, 2"]
    MAR[pos="2, 3"]
    WAF[pos="1, 2"]
    DIV[pos="3, 2"]
    AGE -> DIV
    AGE -> MAR
    MAR -> DIV
```

```
    AGE <- SOU
    SOU -> MAR
    SOU -> WAF
}')

# Plot the DAG
ggdag(dag_6H1, layout = "circle")+
  theme_dag()
```



6H2. Build a series of models to test the implied conditional independencies of the causal graph you used in the previous problem. If any of the tests fail, how do you think the graph needs to be amended? Does the graph need more or fewer arrows? Feel free to nominate variables that aren't in the data.

```
model_south_only <- quap(alist(
    DIV~dnorm(mu , sigma),
    mu<- a + bS * SOU,
    a ~ dnorm( 0, 0.2),
    bS ~ dnorm(0 ,0.5),
    sigma ~ dexp(1)
    ), data = Waffle_data)

precis(model_south_only)
```

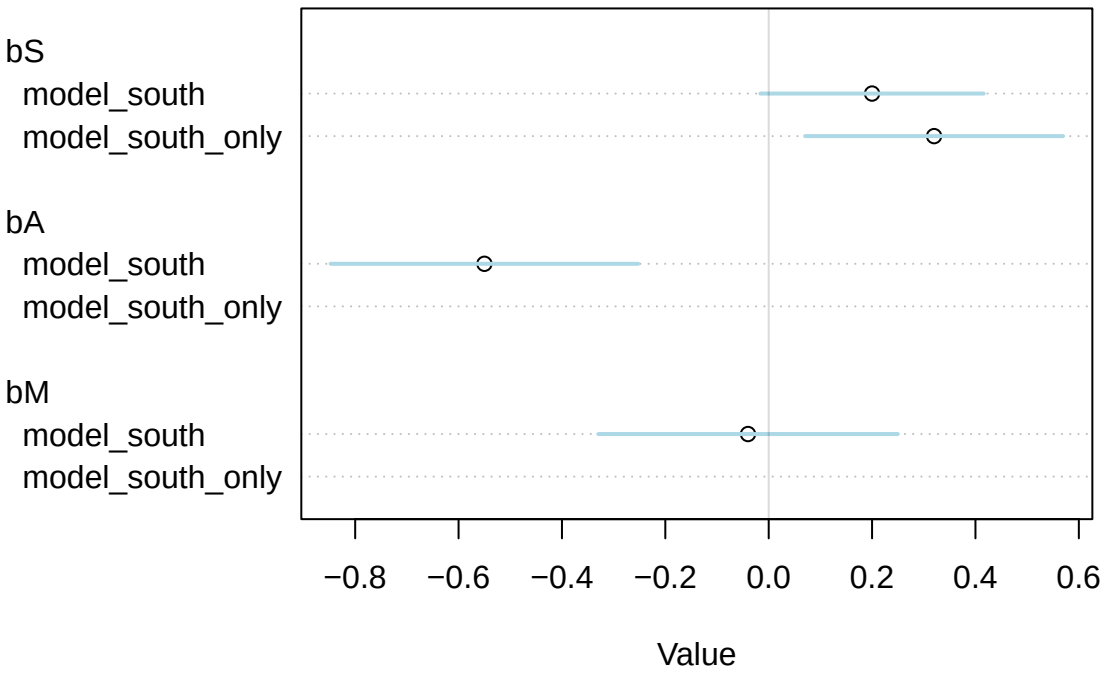
```
##               mean        sd       5.5%     94.5%
## a     1.097477e-05 0.1091383 -0.1744131 0.1744351
```

```
## bS    3.228215e-01 0.1272974  0.1193757 0.5262673
## sigma 9.209274e-01 0.0908938  0.7756616 1.0661933
```

```
model_south <- quap(alist(
    DIV~dnorm(mu , sigma),
    mu<- a + bS * SOU + bA * AGE + bM * MAR,
    a ~ dnorm( 0, 0.2),
    bS ~ dnorm(0 ,0.5),
    bA ~ dnorm(0 ,0.5),
    bM ~ dnorm(0 ,0.5),
    sigma ~ dexp(1)
    ), data = Waffle_data)

precis(model_south)
```

```
##                mean         sd        5.5%      94.5%
## a      1.230597e-05 0.09481550 -0.15152117  0.1515458
## bS     2.029619e-01 0.11020991  0.02682514  0.3790986
## bA    -5.463300e-01 0.15181286 -0.78895622 -0.3037037
## bM    -3.514632e-02 0.14780828 -0.27137250  0.2010798
## sigma  7.614533e-01 0.07545794  0.64085696  0.8820497
```

```
plot(coeftab(model_south_only, model_south),
     pars=c("bS","bA", "bM"),
     col = c("lightblue"))
```



Answer:

When conditioning on age and marriage. The southern states and the divorce rate become conditional independent according to our DAG. When looking at the effect of the southern states we can see that it has become smaller (bs in model_south and in model_south_only). However, it still seems that there is an effect of being married/living in the southern states but this can be due to unobserved confounds.