# Homework 3

Used libraries are:

- rethinking
- ggplot2

### Question 1

Use a model to infer the total causal influence of area on weight. Would increasing the area available to each fox make it heavier (healthier)? You might want to standardize the variables. Regardless, use prior predictive simulation to show that your model's prior predictions stay within the possible outcome range.

There is no need to control on anything on this question. Lets start with standardize the data except the group column.

```
data(foxes)

df <- foxes

d1 <- df %>% scale %>% as.data.frame() %>% mutate(group=df$group)

head(d1)
```

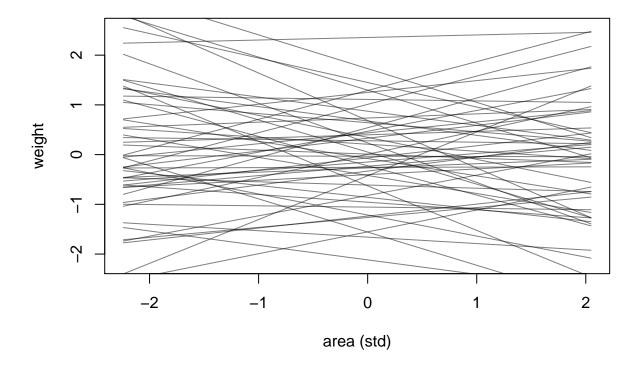
```
##
     group
            avgfood groupsize
                                    area
                                             weight
## 1
         1 -1.924829 -1.524089 -2.239596 0.4141347
## 2
         1 -1.924829 -1.524089 -2.239596 -1.4270464
         2 -1.118035 -1.524089 -1.205508 0.6759540
## 3
## 4
         2 -1.118035 -1.524089 -1.205508 1.3009421
## 5
         3 -1.319734 -1.524089 -1.130106 1.1151348
         3 -1.319734 -1.524089 -1.130106 -1.0807692
```

Lets create the model and check our priors

```
flist <- alist(
  weight ~ dnorm(mu, sigma),
  mu ~ a + b * area,
  a ~ dnorm(0, 1),
  b ~ dnorm(0, 0.5),
  sigma ~ dunif(0, 1)
)
homework_3_question_1 <- quap(flist, data=d1)</pre>
```

Lets visualize the prior models

```
prior <- extract.prior( homework_3_question_1 )
area_seq <- seq( from=min(d1$area) , to=max(d1$area), length.out=50 )
mu <- link( homework_3_question_1 , post=prior, data=list(area= area_seq))
plot( NULL , xlim=range(area_seq) , ylim=range(d1$weight), xlab="area (std)" , ylab="weight" )
for ( i in 1:50 ) lines( area_seq , mu[i,] , col=col.alpha("black",0.5) )</pre>
```



Okay lets check the posterior distribution

```
precis(homework_3_question_1)
```

```
## a mean sd 5.5% 94.5%

## a -8.981161e-07 0.09203683 -0.1470935 0.1470917

## b 1.882896e-02 0.09127046 -0.1270389 0.1646968

## sigma 9.954923e-01 0.06535706 0.8910391 1.0999455
```

No effect found on the dataset of area on weight

## Question 2

Causal flow from food to weight follow two different route, food -> group size -> weight and food -> weight. We focus on both of these effects as we need to know what will happen when we add more food on the territory.

```
flist <- alist(
  weight ~ dnorm(mu, sigma),
  mu ~ a + b * avgfood,
  a ~ dnorm(0, 1),
  b ~ dnorm(0, 1),
  sigma ~ dunif(0, 1)
)

homework_3_question_2 <- quap(flist, data=d1)

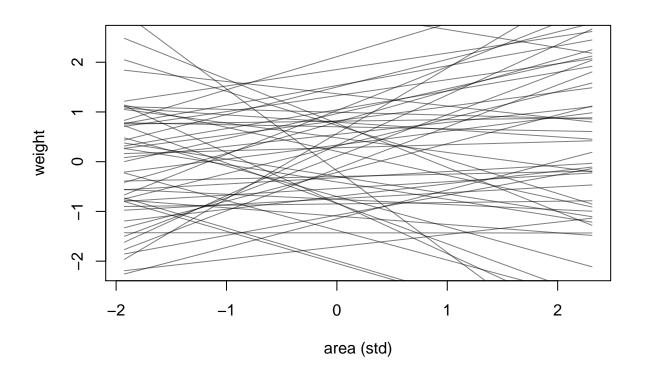
prior <- extract.prior( homework_3_question_1 )

avgfood_seq <- seq( from=min(d1$avgfood) , to=max(d1$avgfood), length.out=50 )

mu <- link( homework_3_question_2 , post=prior, data=list(avgfood= avgfood_seq))

plot( NULL , xlim=range(avgfood_seq) , ylim=range(d1$weight), xlab="area (std)" , ylab="weight" )

for ( i in 1:50 ) lines( avgfood_seq , mu[i,] , col=col.alpha("black",0.5) )</pre>
```



#### precis(homework\_3\_question\_2)

```
## a 4.749883e-07 0.09202551 -0.1470741 0.1470750
## b -2.482507e-02 0.09242136 -0.1725322 0.1228821
## sigma 9.953688e-01 0.06534893 0.8909286 1.0998090
```

# Question 3

We should control for avgfood to find the effect of group\_size on weight, as there is a back door from group size to weight.

There is two path from group size to weight

- group size -> weight
- group size  $\leftarrow$  avg food  $\rightarrow$  weight

```
flist <- alist(
  weight ~ dnorm(mu, sigma),
  mu ~ a + b * avgfood + c * groupsize,
  a ~ dnorm(0, 1),
  b ~ dnorm(0, 0.5),
  c ~ dnorm(0, 0.5),
  sigma ~ dunif(0, 1)
)

homework_3_question_3 <- quap(flist, data=d1)

precis(homework_3_question_3)</pre>
```

```
## mean sd 5.5% 94.5%

## a -3.011148e-07 0.08748754 -0.1398223 0.1398217

## b 4.762362e-01 0.17970128 0.1890389 0.7634336

## c -5.724963e-01 0.17972018 -0.8597239 -0.2852688

## sigma 9.458966e-01 0.06238575 0.8461921 1.0456011
```

So we completed the DAG, the effect of avg food and group\_size and cancel out each other. So we do not see any effect of area.