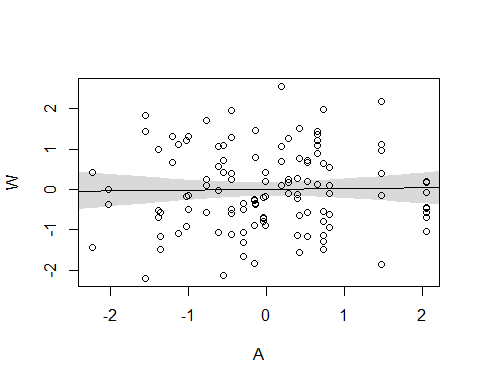
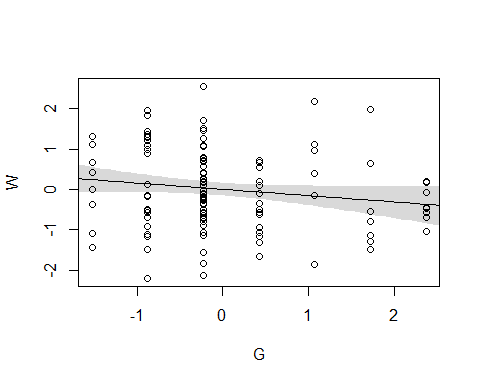
Assignment3\_B04704016林家毅

# Question 1

# Load data  
data(foxes)  
data <- foxes  
  
# Standardize variables  
data$W <- scale( data$weight )  
data$A <- scale( data$area )  
data$G <- scale( data$groupsize )  
data$AF <- scale( data$avgfood )  
  
# Model body weight as a linear function of area  
model.1.1 <- ulam(  
 alist(  
 W ~ dnorm( mu, sigma ),  
 mu <- a + bA\*A,  
 a ~ dnorm( 0, 0.2 ),  
 bA ~ dnorm( 0, 0.5 ),  
 sigma ~ dexp( 1 )  
 ),  
 data=data, chains=4 , cores=4 , iter=1000 )  
  
# Compute posterior probability  
area.seq <- seq( from = -3, to = 3, length.out = 100 )  
area.seq.data <- list( A = area.seq )  
mu <- link( model.1.1 , data = area.seq.data )  
mu.mean <- apply( mu , 2 , mean )  
mu.PI <- apply( mu , 2 , PI , prob = 0.95 )  
  
# Plot results  
plot( W ~ A , data = data )  
lines( area.seq , mu.mean )  
shade( mu.PI , area.seq )



# Model body weight as a linear function of groupsize  
model.1.2 <- ulam(  
 alist(  
 W ~ dnorm( mu , sigma ),  
 mu <- a + bG\*G,  
 a ~ dnorm( 0, 0.2 ),  
 bG ~ dnorm( 0, 0.5 ),  
 sigma ~ dexp( 1 )  
 ),  
 data=data, chains=4 , cores=4 , iter=1000 )  
  
# Compute posterior probability  
groupsize.seq <- seq( from = -3 , to = 3 , length.out = 100 )  
groupsize.seq.data <- list ( G = groupsize.seq )  
mu <- link( model.1.2 , data = groupsize.seq.data )  
mu.mean <- apply( mu , 2 , mean )  
mu.PI <- apply( mu, 2 , PI , prob = 0.95 )  
  
# Plot results  
plot( W ~ G , data = data )  
lines( groupsize.seq , mu.mean )  
shade( mu.PI , groupsize.seq )



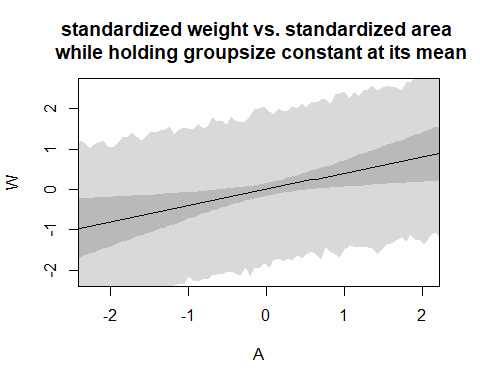
# Conclusion:  
# According to the plots, it seems that both models are not great for predicting body weight of foxes.

# Question 2

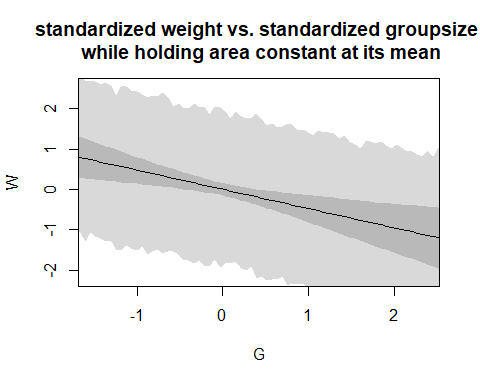
# Fit the multiple linear regression model with weight as the outcome and both area and groupsize as predictor variables.  
model.2 <- ulam(  
 alist(  
 W ~ dnorm( mu, sigma ),  
 mu <- a + bG\*G + bA\*A,  
 a ~ dnorm( 0, 0.2 ),  
 bG ~ dnorm( 0, 0.5 ),  
 bA ~ dnorm( 0, 0.5 ),  
 sigma ~ dexp( 1 )  
 ),  
 data=data, chains=4 , cores=4 , iter=1000 )  
  
precis(model.2)

## mean sd 5.5% 94.5% n\_eff Rhat  
## a 0.0009455506 0.07975736 -0.1217046 0.1289047 1393.713 1.001920  
## bG -0.4758110573 0.15136926 -0.7180372 -0.2290681 1117.050 1.005357  
## bA 0.4031645192 0.14919092 0.1538559 0.6374846 1154.750 1.004174  
## sigma 0.9621071923 0.06394003 0.8665883 1.0685038 1500.743 1.002149

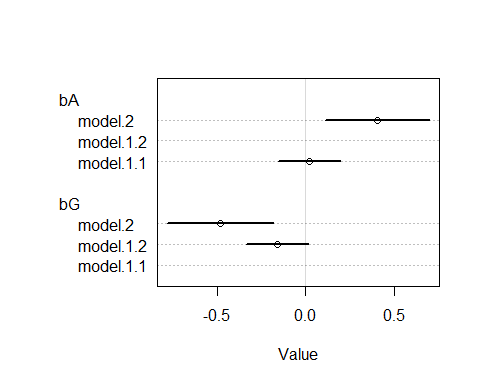
# Plot territory size vs. weight, holding groupsize constant at its mean  
area.predict.data <- data.frame( A = area.seq , G = 0 )  
  
mu <- link( model.2 , data = area.predict.data )  
mu.mean <- apply( mu, 2 , mean )  
mu.PI <- apply( mu , 2 , PI , prob=0.95 )  
pred.sim <- sim( model.2 , data=area.predict.data , n=1e4 )  
pred.PI <- apply( pred.sim , 2 , PI, prob=0.95 )  
  
plot( W ~ A , data = data , type = "n" )  
lines( area.seq , mu.mean )  
shade( mu.PI , area.seq )  
shade( pred.PI , area.seq )  
title('standardized weight vs. standardized area \n while holding groupsize constant at its mean')



# Plot groupsize vs. weight, holding area constant at its mean  
groupsize.predict.data <- data.frame( A = 0 , G = groupsize.seq )  
  
mu <- link( model.2 , groupsize.predict.data )  
mu.mean <- apply( mu , 2 , mean )  
mu.PI <- apply( mu , 2 , PI , prob = 0.95 )  
pred.sim <- sim( model.2 , data=groupsize.predict.data , n=1e4 )  
pred.PI <- apply( pred.sim , 2 , PI, prob=0.95 )  
  
plot( W ~ G , data = data , type = "n" )  
lines( area.seq , mu.mean )  
shade( mu.PI , area.seq )  
shade( pred.PI , area.seq )  
title('standardized weight vs. standardized groupsize \n while holding area constant at its mean')



# Compare the coefficients of each model  
plot( coeftab( model.1.1 , model.1.2 , model.2 ), par=c( "bA" , "bG" ) )



# Conclusion:  
# This model implies that if the area is bigger, then the body weights of foxes would be larger. On the other hand, if the groupsize is bigger, the foxes would have less weight. The differece between the models of two questions comes from different linear regression modeling approach. From the results, we could find out that both variables together are important for predicting wieght of foxes.

# Question 3

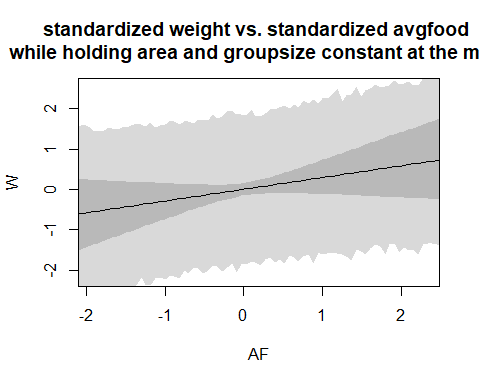
# Model weight as an additive function of avgfood and groupsize  
model.3.1 <- ulam(  
 alist(  
 W ~ dnorm( mu, sigma ),  
 mu <- a + bAF\*AF + bG\*G,  
 a ~ dnorm( 0, 0.2 ),  
 bAF ~ dnorm( 0, 0.5 ),  
 bG ~ dnorm( 0, 0.5 ),  
 sigma ~ dexp( 1 )  
 ),  
 data=data, chains=4 , cores=4 , iter=1000 )  
  
precis(model.3.1)

## mean sd 5.5% 94.5% n\_eff Rhat  
## a -0.001298409 0.08229227 -0.1320918 0.1260591 1678.7899 0.9989141  
## bAF 0.468299387 0.18796633 0.1500930 0.7531573 987.7618 1.0055052  
## bG -0.567924041 0.18334117 -0.8573892 -0.2728832 973.0607 1.0065641  
## sigma 0.963343336 0.06545838 0.8653232 1.0732129 1559.6390 0.9991085

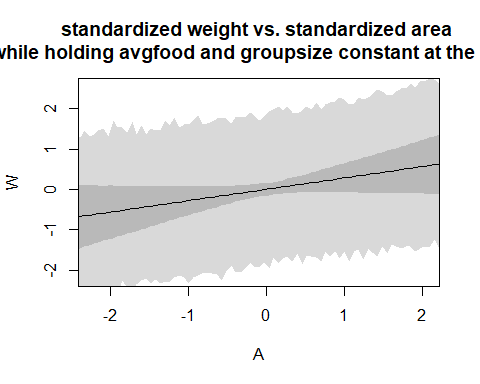
# Model weight as an additive function of avgfood, groupsize and area  
model.3.2 <- ulam(  
 alist(  
 W ~ dnorm( mu, sigma ),  
 mu <- a + bAF\*AF + bG\*G + bA\*A,  
 a ~ dnorm( 0, 0.2 ),  
 bAF ~ dnorm( 0, 0.5 ),  
 bG ~ dnorm( 0, 0.5 ),  
 bA ~ dnorm( 0, 0.5 ),  
 sigma ~ dexp( 1 )  
 ),  
 data=data, chains=4 , cores=4 , iter=1000 )  
  
precis(model.3.2)

## mean sd 5.5% 94.5% n\_eff Rhat  
## a 0.002270921 0.07977705 -0.12673910 0.1304339 1746.824 1.0005824  
## bAF 0.291414814 0.20950023 -0.03632898 0.6226812 1275.681 0.9991028  
## bG -0.634910827 0.19214302 -0.94727809 -0.3402040 1467.967 1.0003584  
## bA 0.281467898 0.16486469 0.02213315 0.5474372 1617.587 0.9991858  
## sigma 0.955596804 0.06587960 0.85425613 1.0658302 1650.133 1.0004250

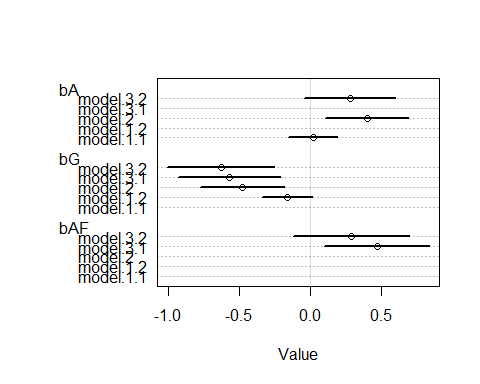
# Plot weight as a function of avgfood, holding groupsize and area at their respective means  
avgfood.seq <- seq( from = -3 , to = 3 , length.out = 100 )  
avgfood.predict.data <- data.frame(  
 A = 0,  
 G = 0,  
 AF = avgfood.seq  
)  
  
mu <- link( model.3.2 , data = avgfood.predict.data )  
mu.mean <- apply( mu, 2 , mean )  
mu.PI <- apply( mu , 2 , PI , prob=0.95 )  
  
pred.sim <- sim( model.3.2 , data = avgfood.predict.data , n = 1e4 )  
pred.PI <- apply( pred.sim , 2 , PI, prob=0.95 )  
  
plot( W ~ AF , data = data , type = "n" )  
lines( avgfood.seq , mu.mean )  
shade( mu.PI , avgfood.seq )  
shade( pred.PI , avgfood.seq )  
title('standardized weight vs. standardized avgfood \n while holding area and groupsize constant at the mean')



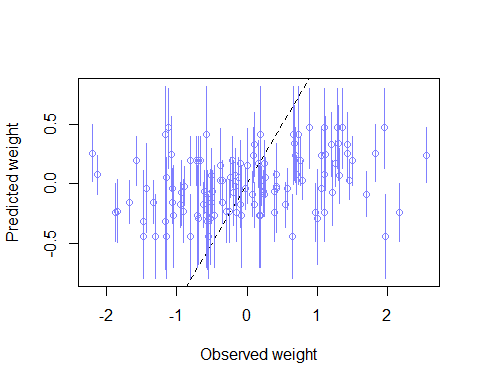
# Plot weight as a function of area, holding groupsize and avgfood at their respective means  
area.seq <- seq( from = -3 , to = 3 , length.out = 100 )  
area.predict.data <- data.frame(  
 A = area.seq,  
 G = 0,  
 AF = 0  
)  
  
mu <- link( model.3.2 , data = area.predict.data )  
mu.mean <- apply( mu, 2 , mean )  
mu.PI <- apply( mu , 2 , PI , prob=0.95 )  
  
pred.sim <- sim( model.3.2 , data = area.predict.data , n = 1e4 )  
pred.PI <- apply( pred.sim , 2 , PI, prob=0.95 )  
  
plot( W ~ A , data = data , type = "n" )  
lines( area.seq , mu.mean )  
shade( mu.PI , area.seq )  
shade( pred.PI , area.seq )  
title('standardized weight vs. standardized area \n while holding avgfood and groupsize constant at the mean')



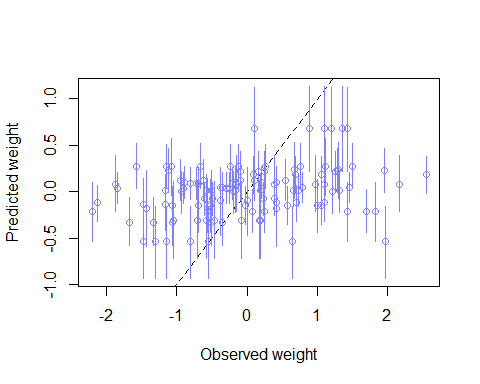
# Compare the coefficients of each model  
plot( coeftab( model.1.1 , model.1.2 , model.2 , model.3.1 , model.3.2 ), par=c( "bA" , "bG" , "bAF" ) )



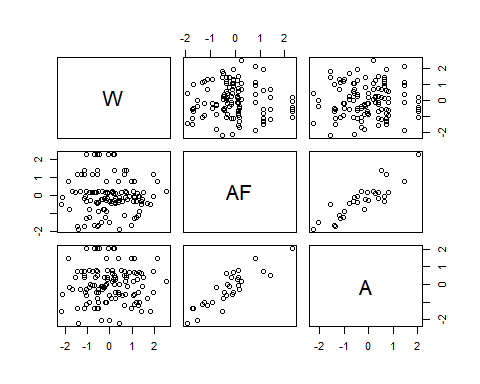
# Inspect predicting results of the model that uses avgfood and groupsize as predictor variables  
mu <- link( model.3.1 )  
mu.mean <- apply( mu , 2 , mean )  
mu.PI <- apply( mu , 2 , PI , prob=0.95 )  
pre.sim <- sim( model.3.1 , n=1e4 )  
pre.PI <- apply( pre.sim , 2 , PI , prob=0.95 )  
plot( mu.mean ~ data$W , col=rangi2 , ylim=range(mu.PI) , xlab="Observed weight" , ylab="Predicted weight" )  
abline( a=0 , b=1 , lty=2 )  
for ( i in 1:nrow(data) ) lines( rep(data$W[i],2) , mu.PI[,i] , col=rangi2 )



# Inspect predicting results of the model that uses area and groupsize as predictor variables  
mu <- link( model.2 )  
mu.mean <- apply( mu , 2 , mean )  
mu.PI <- apply( mu , 2 , PI , prob=0.95 )  
pre.sim <- sim( model.2 , n=1e4 )  
pre.PI <- apply( pre.sim , 2 , PI , prob=0.95 )  
plot( mu.mean ~ data$W , col=rangi2 , ylim=range(mu.PI) , xlab="Observed weight" , ylab="Predicted weight" )  
abline( a=0 , b=1 , lty=2 )  
for ( i in 1:nrow(data) ) lines( rep(data$W[i],2) , mu.PI[,i] , col=rangi2 )



# (a)  
# By comparing the simulating results of model.3.1 and model.2, it shows that model.2 performs better. Thus, we could consider area to be the better predictor variable of body weight.  
  
pairs( ~ W + AF + A , data )



# (b)  
# In model.3.2, since both of the predictor variables, area and avgfood, are positively correlated with one another. In addition, both variables are positively correlated with the outcome. Thus, if we put these two variables in the same model, their effects would be reduced, and the standard error would be larger.