HW 4

Nathan Hawkins

2/11/2021

```
vo2 <- read.table("vo2.dat", header = TRUE)
head(vo2)
```

```
    ID Gender Age1   BMI MPH  HR RPE MaxVO2ML
 1   1      0   12 16.49 4.6 178  14    43.69
 2  95      0   12 17.80 6.0 171  10    48.07
 3 102      0   12 20.48 4.8 185  13    45.73
 4 111      0   12 16.46 5.4 155  13    46.65
 5 120      0   12 15.94 4.7 154  14    55.56
 6 121      0   12 17.51 5.0 172  14    50.00
```

1

Running MLR with the 6 covariates, I get a DIC of 756.6.

```
mdl <- "
  model {

  for(i in 1:120){
    y[i] ~ dnorm(mu[i], prec)
    mu[i] <- b0 +  bgen*gender[i] + bbmi*bmi[i] + brpe*rpe[i]+ bhr*hr[i] +
bmph*mph[i] + bage*age[i]
  }

  b0 ~ dnorm(0 , 0.001)
  bgen ~ dnorm(0, 0.001)
  bbmi ~ dnorm(0, 0.001)
  brpe ~ dnorm(0, 0.001)
  bhr ~ dnorm(0, 0.001)
  bmph ~ dnorm(0, 0.001)
  bage ~ dnorm(0, 0.001)

  prec <- 1/vv
  vv ~ dgamma(1.5, 0.2)

  }

"

mph <- vo2$MPH
```

```r
bmi <- vo2$BMI
gender <- vo2$Gender
y <- vo2$MaxVO2ML
rpe <- vo2$RPE
hr <- vo2$HR
age <- vo2$Age1


# Standardize the covariates
mph <- (mph-mean(mph))/sd(mph)
bmi <- (bmi-mean(bmi))/sd(bmi)
gender <- (gender-mean(gender))/sd(gender)
rpe <- (rpe-mean(rpe))/sd(rpe)
age <- (age-mean(age))/sd(age)
hr <- (hr-mean(hr))/sd(hr)
mphhr <- mph*hr




writeLines(mdl, 'vo2.txt')

#What to use
data.jags <- c('y', 'gender', 'bmi', 'rpe', 'mph', 'hr', 'age')

#What to call it in the output
parms.jags <- c('b0', "brpe", 'bgen', 'bbmi', 'bmph', 'vv', 'bhr', 'bage')

vo2.sim <- jags(data= data.jags, parameters.to.save = parms.jags,
                  model.file = 'vo2.txt', inits = NULL,
                  n.iter = 6000, n.thin = 1, n.chains = 4,
                  n.burnin = 1000)
```

```
 module glm loaded

 Compiling model graph
    Resolving undeclared variables
    Allocating nodes
 Graph information:
    Observed stochastic nodes: 120
    Unobserved stochastic nodes: 8
    Total graph size: 1183

 Initializing model
```

```r
vo2.sim
```

```
 Inference for Bugs model at "vo2.txt", fit using jags,
  4 chains, each with 6000 iterations (first 1000 discarded)
```

```
  n.sims = 20000 iterations saved
          mu.vect sd.vect    2.5%     25%     50%     75%   97.5%  Rhat n.eff
 b0        48.773   0.485  47.818  48.443  48.777  49.100  49.720 1.001 17000
 bage       0.483   0.557  -0.604   0.109   0.483   0.864   1.561 1.001 13000
 bbmi      -3.181   0.523  -4.198  -3.534  -3.186  -2.830  -2.154 1.001 20000
 bgen       2.690   0.561   1.587   2.316   2.690   3.064   3.807 1.001 20000
 bhr       -1.599   0.512  -2.612  -1.938  -1.599  -1.256  -0.590 1.001 20000
 bmph       3.802   0.583   2.658   3.408   3.803   4.194   4.957 1.001 20000
 brpe      -1.122   0.492  -2.082  -1.453  -1.124  -0.793  -0.140 1.001 20000
 vv        28.195   3.561  22.038  25.725  27.893  30.389  36.048 1.001  9600
 deviance 748.584   4.003 742.758 745.646 747.955 750.788 758.213 1.001  7900

 For each parameter, n.eff is a crude measure of effective sample size,
 and Rhat is the potential scale reduction factor (at convergence, Rhat=1).

 DIC info (using the rule, pD = var(deviance)/2)
 pD = 8.0 and DIC = 756.6
 DIC is an estimate of expected predictive error (lower deviance is better).
```

2

Running the same model but without age, I get a DIC of 755.1.

```
library(R2jags)
mdl2 <- "
  model {

  for(i in 1:120){
    y[i] ~ dnorm(mu[i], prec)
    mu[i] <- b0 +  bgen*gender[i] + bbmi*bmi[i] + brpe*rpe[i]+ bhr*hr[i] +
bmph*mph[i]
  }

  b0 ~ dnorm(0 , 0.001)
  bgen ~ dnorm(0, 0.001)
  bbmi ~ dnorm(0, 0.001)
  brpe ~ dnorm(0, 0.001)
  bhr ~ dnorm(0, 0.001)
  bmph ~ dnorm(0, 0.001)

  prec <- 1/vv
  vv ~ dgamma(1.5, 0.2)

  }

"
# Assign simple names
mph <- vo2$MPH
```

```r
bmi <- vo2$BMI
gender <- vo2$Gender
y <- vo2$MaxVO2ML
rpe <- vo2$RPE
hr <- vo2$HR
age <- vo2$Age1


# Standardize the covariates
mph <- (mph-mean(mph))/sd(mph)
bmi <- (bmi-mean(bmi))/sd(bmi)
gender <- (gender-mean(gender))/sd(gender)
rpe <- (rpe-mean(rpe))/sd(rpe)
age <- (age-mean(age))/sd(age)
hr <- (hr-mean(hr))/sd(hr)
```

```r
writeLines(mdl2, 'vo2_v2.txt')

#What to use
data.jags <- c('y', 'gender', 'bmi', 'rpe', 'mph', 'hr')

#What to call it in the output
parms.jags <- c('b0', "brpe", 'bgen', 'bbmi', 'bmph', 'vv', 'bhr')

vo2_v2.sim <- jags(data= data.jags, parameters.to.save = parms.jags,
                   model.file = 'vo2_v2.txt', inits = NULL,
                   n.iter = 6000, n.thin = 1, n.chains = 4,
                   n.burnin = 1000)
```

```
 Compiling model graph
    Resolving undeclared variables
    Allocating nodes
 Graph information:
    Observed stochastic nodes: 120
    Unobserved stochastic nodes: 7
    Total graph size: 1056

 Initializing model
```

```r
vo2_v2.sim
```

```
 Inference for Bugs model at "vo2_v2.txt", fit using jags,
  4 chains, each with 6000 iterations (first 1000 discarded)
  n.sims = 20000 iterations saved
```

```
          mu.vect sd.vect    2.5%     25%     50%     75%   97.5%  Rhat n.eff
b0         48.773   0.483  47.830  48.445  48.770  49.099  49.715 1.001 20000
bbmi       -3.041   0.500  -4.025  -3.380  -3.042  -2.704  -2.062 1.001 13000
bgen        2.579   0.542   1.510   2.223   2.581   2.944   3.639 1.001 12000
bhr        -1.677   0.501  -2.661  -2.015  -1.674  -1.338  -0.690 1.001  5200
bmph        3.997   0.536   2.953   3.636   3.996   4.351   5.051 1.001 20000
brpe       -1.105   0.492  -2.064  -1.434  -1.108  -0.775  -0.134 1.001 20000
vv         28.070   3.462  22.049  25.589  27.801  30.288  35.516 1.001  4500
deviance  748.315   3.700 743.031 745.605 747.671 750.328 757.167 1.001 20000

For each parameter, n.eff is a crude measure of effective sample size,
and Rhat is the potential scale reduction factor (at convergence, Rhat=1).

DIC info (using the rule, pD = var(deviance)/2)
pD = 6.8 and DIC = 755.2
DIC is an estimate of expected predictive error (lower deviance is better).
```

3

I would choose the second model because it has a lower DIC so it is better at predicting with fewer covariates.

4

Same model but with mph*hr. This model has a DIC of 753.5.

```
library(R2jags)
mdl3 <- "
  model {

  for(i in 1:120){
    y[i] ~ dnorm(mu[i], prec)
    mu[i] <- b0 +  bgen*gender[i] + bbmi*bmi[i] + brpe*rpe[i]+ bhr*hr[i] +
bmph*mph[i] + bmphhr*mphhr[i]
  }

  b0 ~ dnorm(0 , 0.001)
  bgen ~ dnorm(0, 0.001)
  bbmi ~ dnorm(0, 0.001)
  brpe ~ dnorm(0, 0.001)
  bhr ~ dnorm(0, 0.001)
  bmph ~ dnorm(0, 0.001)
  bmphhr ~ dnorm(0, .001)

  prec <- 1/vv
  vv ~ dgamma(1.5, 0.2)
```

```
  }

"
#Assigning simple names for covariates
mph <- vo2$MPH
bmi <- vo2$BMI
gender <- vo2$Gender
y <- vo2$MaxVO2ML
rpe <- vo2$RPE
hr <- vo2$HR
age <- vo2$Age1


# Standardize the covariates
mph <- (mph-mean(mph))/sd(mph)
bmi <- (bmi-mean(bmi))/sd(bmi)
gender <- (gender-mean(gender))/sd(gender)
rpe <- (rpe-mean(rpe))/sd(rpe)
age <- (age-mean(age))/sd(age)
hr <- (hr-mean(hr))/sd(hr)
mphhr <- mph*hr
mphhr = (mphhr - mean(mphhr))/sd(mphhr)




writeLines(mdl3, 'vo2_v3.txt')

#What to use
data.jags <- c('y', 'gender', 'bmi', 'rpe', 'mph', 'hr', 'mphhr')

#What to call it in the output
parms.jags <- c('b0', "brpe", 'bgen', 'bbmi', 'bmph', 'vv', 'bhr', 'bmphhr')

vo2_v3.sim <- jags(data= data.jags, parameters.to.save = parms.jags,
                   model.file = 'vo2_v3.txt', inits = NULL,
                   n.iter = 6000, n.thin = 1, n.chains = 4,
                   n.burnin = 1000)

 Compiling model graph
    Resolving undeclared variables
    Allocating nodes
 Graph information:
    Observed stochastic nodes: 120
    Unobserved stochastic nodes: 8
    Total graph size: 1293

 Initializing model
```

```
vo2_v3.sim

 Inference for Bugs model at "vo2_v3.txt", fit using jags,
  4 chains, each with 6000 iterations (first 1000 discarded)
  n.sims = 20000 iterations saved
          mu.vect sd.vect    2.5%     25%     50%     75%   97.5%  Rhat n.eff
 b0        48.780   0.476  47.854  48.460  48.776  49.100  49.709 1.001  9100
 bbmi      -3.236   0.506  -4.229  -3.575  -3.239  -2.897  -2.251 1.001  8000
 bgen       2.714   0.549   1.639   2.344   2.714   3.080   3.791 1.001 20000
 bhr       -1.771   0.505  -2.750  -2.112  -1.771  -1.431  -0.783 1.001  5000
 bmph       3.556   0.576   2.418   3.171   3.561   3.938   4.691 1.001 20000
 bmphhr    -1.025   0.542  -2.075  -1.390  -1.026  -0.665   0.043 1.001 20000
 brpe      -1.226   0.491  -2.182  -1.553  -1.231  -0.898  -0.258 1.001 20000
 vv        27.554   3.422  21.576  25.158  27.296  29.681  34.910 1.001 20000
 deviance 745.713   3.986 739.932 742.800 745.047 747.894 755.149 1.001 20000

 For each parameter, n.eff is a crude measure of effective sample size,
 and Rhat is the potential scale reduction factor (at convergence, Rhat=1).

 DIC info (using the rule, pD = var(deviance)/2)
 pD = 7.9 and DIC = 753.7
 DIC is an estimate of expected predictive error (lower deviance is better).
```

5

I prefer the last model because it has the lowest DIC. Adding the interaction term is worth it.

6 SAS code works, doesn't knit.

library(SASmarkdown)

saspath <- "C:/Program Files/SASHome/SASFoundation/9.4/sas.exe" sasopts <- "-nosplash -ls 75" knitr::opts_chunk$set(engine="sas", engine.path=saspath, engine.opts=sasopts, comment=NA)

knitr::opts_chunk$get()$engine knitr::opts_chunk$get()$engine.path knitr::opts_chunk$get()$engine.opts

1 Sas

DIC is 756.5

data vo2; infile 'C:/Users/nateh/Documents/Stat 451/vo2.dat'; input ID Gender Age BMI MPH HR RPE MAXVO2ML; run;

proc standard data=vo2 out=vo2 mean=0 std=1; var ID Gender Age BMI MPH HR RPE:; run;

proc mcmc data = vo2 nbi = 30000 nmc = 300000 thin = 30 outpost = 'C:/Users/nateh/Documents/Stat 451/MLRHW4.sas7bdat' dic propcov=quanew monitor=(*parms*) stats = all; parms b0 0; parms bmph 0; parms bage 0; parms bhr 0; parms brpe 0; parms bgen 0; parms bbmi 0; parms vv 1.5; prior b0 ~ normal(0, var = 1000); prior bmph: ~ normal(0, var = 1000); prior bage: ~ normal(0, var = 1000); prior bhr: ~ normal(0, var = 1000); prior brpe: ~ normal(0, var = 1000); prior bbmi: ~ normal(0, var = 1000); prior bgen: ~ normal(0, var = 1000); prior vv ~ gamma(1.5, scale = 5); mu = b0 + bgen*Gender + bbmi*BMI + brpe*RPE + bhr*HR + bmph*MPH + bage*Age; model MAXVO2ML ~ normal(mu, var = vv); run;

2

Model without age, DIC is 755.1

data vo2; infile 'C:/Users/nateh/Documents/Stat 451/vo2.dat'; input ID Gender Age BMI MPH HR RPE MAXVO2ML; run;

proc standard data=vo2 out=vo2 mean=0 std=1; var ID Gender Age BMI MPH HR RPE:; run;

proc mcmc data = vo2 nbi = 30000 nmc = 300000 thin = 30 outpost = 'C:/Users/nateh/Documents/Stat 451/MLRHW4_2.sas7bdat' dic propcov=quanew monitor=(*parms*); parms b0 0; parms bmph 0; parms bhr 0; parms brpe 0; parms bgen 0; parms bbmi 0; parms vv 1.5; prior b0 ~ normal(0, var = 1000); prior bmph: ~ normal(0, var = 1000); prior bhr: ~ normal(0, var = 1000); prior brpe: ~ normal(0, var = 1000); prior bbmi: ~ normal(0, var = 1000); prior bgen: ~ normal(0, var = 1000); prior vv ~ gamma(1.5, scale = 5); mu = b0 + bgen*Gender + bbmi*BMI + brpe*RPE + bhr*HR + bmph*MPH; model MAXVO2ML ~ normal(mu, var = vv); run;

3 Sas

2nd model is better, lower DIC

4 SAS

Model without age, with mphhr. DIC is 753.6

data vo2; infile 'C:/Users/nateh/Documents/Stat 451/vo2.dat'; input ID Gender Age BMI MPH HR RPE MAXVO2ML; run;

Standardize Variables

proc standard data=vo2 out=vo2 mean=0 std=1; var ID Gender Age BMI MPH HR RPE:; run;

proc mcmc data = vo2 nbi = 30000 nmc = 300000 thin = 30 outpost = 'C:/Users/nateh/Documents/Stat 451/MLRHW4_3.sas7bdat' dic propcov=quanew monitor=(*parms*); parms b0 0; parms bmph 0; parms bhr 0; parms brpe 0; parms bgen 0;

parms bbmi 0; parms binteraction 0; parms vv 1.5; prior b0 ~ normal(0, var = 1000); prior bmph: ~ normal(0, var = 1000); prior bhr: ~ normal(0, var = 1000); prior brpe: ~ normal(0, var = 1000); prior bbmi: ~ normal(0, var = 1000); prior bgen: ~ normal(0, var = 1000); prior binteraction ~ normal(0, var = 1000); prior vv ~ gamma(1.5, scale = 5); mu = b0 + bgen*Gender + bbmi*BMI + brpe*RPE + bhr*HR + bmph*MPH + binteraction*(MPH*HR); model MAXVO2ML ~ normal(mu, var = vv); run;

5 SAS

The results are the same, the third model is the best.