STAT 431 Homework 6 Solutions

Robert Garrett

Problem 1

Part a

Data and initial values:

```
lifeexpdiff = read.csv("lifeexpdiff.csv", row.names=1)
years = c(1960,1970,1980,1990,2000,2010)
d = list(Y = lifeexpdiff,
         X = years,
         Xbar = mean(years))
inits = list(list(tausq.y=1, beta1=0, beta2=0,
                  sigma.alpha1=1, sigma.alpha2=1),
             list(tausq.y=100, beta1=100, beta2=100,
                  sigma.alpha1=0.1, sigma.alpha2=0.1),
             list(tausq.y=0.01, beta1=-100, beta2=-100,
                  sigma.alpha1=10, sigma.alpha2=10))
```

JAGS model code:

```
data {
  dim.Y <- dim(Y)
}

model {
  for(i in 1:dim.Y[1]) {

    for(j in 1:dim.Y[2]) {
      Y[i,j] ~ dnorm(mu[i,j], tausq.y)
      mu[i,j] <- alpha[i,1] + alpha[i,2] * (X[j] - Xbar)
    }

    alpha[i,1] ~ dnorm(beta1, 1 / sigma.alpha1^2)
    alpha[i,2] ~ dnorm(beta2, 1 / sigma.alpha2^2)
  }

  tausq.y ~ dgamma(0.001, 0.001)
  sigma.y <- 1 / sqrt(tausq.y)

  beta1 ~ dnorm(0.0, 1.0E-6)
  beta2 ~ dnorm(0.0, 1.0E-6)
  sigma.alpha1 ~ dexp(0.001)
  sigma.alpha2 ~ dexp(0.001)
}
```

Part b

```
m = jags.model("prob1.bug", d, inits, n.chains=3)

#Burn in run
x = coda.samples(m, c("beta1","beta2","sigma.y","sigma.alpha1","sigma.alpha2"),
                 n.iter=10000)
#Assess convergence here (no need to include in solution)

#Draw final posterior samples
x = coda.samples(m, c("beta1","beta2","sigma.y","sigma.alpha1","sigma.alpha2"),
                 n.iter=100000)
```

```
summary(x)
```

```
##
## Iterations = 11001:111000
## Thinning interval = 1
## Number of chains = 3
## Sample size per chain = 1e+05
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##                   Mean       SD  Naive SE Time-series SE
## beta1         0.116751 0.190402 3.476e-04      3.519e-04
## beta2        -0.005432 0.003688 6.733e-06      7.651e-06
## sigma.alpha1  1.328738 0.140884 2.572e-04      3.636e-04
## sigma.alpha2  0.024229 0.002894 5.283e-06      8.646e-06
## sigma.y       0.382594 0.019252 3.515e-05      5.028e-05
##
## 2. Quantiles for each variable:
##
##                  2.5%       25%       50%      75%    97.5%
## beta1        -0.25709 -0.010168  0.116447  0.24367 0.490750
## beta2        -0.01269 -0.007885 -0.005425 -0.00299 0.001841
## sigma.alpha1  1.08646  1.229669  1.317160  1.41542 1.636149
## sigma.alpha2  0.01917  0.022202  0.024011  0.02603 0.030522
## sigma.y       0.34719  0.369276  0.381768  0.39504 0.422770
```

Part c

Each of the specified credible intervals contain 0. Since the values are centered around the US average, and we are comparing every US state, the values themselves should be centered around 0. So, for intercept β_1, the average differenced life expectancy at time 0 (1995) should be around 0. Also, for the slope β_2, the average differenced life expectancy should not change over time, as the values are centered around the US average at each observed year.

Problem 2

Part a

Data and initial values:

```r
d <- list(Y = lifeexpdiff,
          X = years,
          Xbar = mean(years),
          Omega0 = rbind(c(1, 0),
                         c(0, 0.0005)),
          nu = 2,
          mu0 = c(0,0),
          Sigma0.inv = rbind(c(1.0E-6, 0),
                             c(0, 1.0E-6)))
inits <- list(list(tausq.y=1, beta=c(0,0),
                   Omega.inv=diag(2)),
              list(tausq.y=100, beta=c(100,100),
                   Omega.inv=100*diag(2)),
              list(tausq.y=0.01, beta=c(-100,-100),
                   Omega.inv=0.01*diag(2)))
```

JAGS model code:

```
data {
  dim.Y <- dim(Y)
}

model {
  for(i in 1:dim.Y[1]) {

    for(j in 1:dim.Y[2]) {
      Y[i,j] ~ dnorm(mu[i,j], tausq.y)
      mu[i,j] <- alpha[i,1] + alpha[i,2] * (X[j] - Xbar)
    }

    alpha[i,1:2] ~ dmnorm(beta, Omega.inv)
  }

  tausq.y ~ dgamma(0.001, 0.001)
  sigma.y <- 1 / sqrt(tausq.y)

  beta ~ dmnorm(mu0, Sigma0.inv)
  Omega.inv ~ dwish(nu*Omega0, nu)
  Omega <- inverse(Omega.inv)

  rho <- Omega[1,2] / sqrt(Omega[1,1] * Omega[2,2])
  rho_greater_0 <- rho > 0
}
```

Part b

```r
m <- jags.model("prob2.bug", d, inits, n.chains=3)

#Burn in run
x = coda.samples(m, c("beta","sigma.y","Omega","rho","rho_greater_0"),
```

```r
                 n.iter=10000)
#Assess convergence here (no need to include in solution)

#Draw final posterior samples
x = coda.samples(m, c("beta","sigma.y","Omega","rho","rho_greater_0"),
                 n.iter=100000)
```

```r
summary(x)
```

```
##
## Iterations = 10001:110000
## Thinning interval = 1
## Number of chains = 3
## Sample size per chain = 1e+05
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##                    Mean        SD  Naive SE Time-series SE
## Omega[1,1]     1.7526041 0.3703332 6.761e-04      7.052e-04
## Omega[2,1]     0.0056112 0.0050997 9.311e-06      1.078e-05
## Omega[1,2]     0.0056112 0.0050997 9.311e-06      1.078e-05
## Omega[2,2]     0.0005865 0.0001396 2.549e-07      3.273e-07
## beta[1]        0.1167458 0.1886659 3.445e-04      3.521e-04
## beta[2]       -0.0054382 0.0036715 6.703e-06      7.573e-06
## rho            0.1736895 0.1465730 2.676e-04      3.146e-04
## rho_greater_0  0.8779067 0.3273941 5.977e-04      6.474e-04
## sigma.y        0.3828238 0.0193541 3.534e-05      5.036e-05
##
## 2. Quantiles for each variable:
##
##                    2.5%        25%        50%        75%     97.5%
## Omega[1,1]     1.1708360  1.4901183  1.7040146  1.9615238 2.6094023
## Omega[2,1]    -0.0039393  0.0022745  0.0053851  0.0087195 0.0163393
## Omega[1,2]    -0.0039393  0.0022745  0.0053851  0.0087195 0.0163393
## Omega[2,2]     0.0003669  0.0004878  0.0005684  0.0006652 0.0009111
## beta[1]       -0.2547722 -0.0090186  0.1169758  0.2421320 0.4892334
## beta[2]       -0.0126759 -0.0078823 -0.0054328 -0.0029879 0.0017820
## rho           -0.1234506  0.0758547  0.1777503  0.2757169 0.4488084
## rho_greater_0  0.0000000  1.0000000  1.0000000  1.0000000 1.0000000
## sigma.y        0.3472312  0.3693745  0.3819875  0.3953664 0.4231258
```

Part c

From the summary table in part b, we see the average value of the binary variable "rho_greater_0" is 0.8779. So, the posterior probability that $\rho > 0$ is 0.8779.

Problem 3

Part (a) (i)

$$
\begin{aligned}
0 = Prob(Y_1 = 1, Y_2 = 2) &\Rightarrow 0 = \sum_x Prob(Y_1 = 1, Y_2 = 1 | X = x) Prob(X = x) && (1)\\
&\Rightarrow 0 = \sum_x Prob(Y_1 = 1 | X = x) Prob(Y_2 = 1 | X = x) Prob(X = x) && (2)\\
&\Rightarrow 0 = \sum_x g(x)^2 Prob(X = x) && (3)\\
&\Rightarrow 0 = \sum_{\{x: Prob(X=x)=0\}} g(x)^2 Prob(X = x) + \sum_{\{x: Prob(X=x)>0\}} g(x)^2 Prob(X = x) && (4)\\
&\Rightarrow 0 = \sum_{\{x: Prob(X=x)=0\}} g(x)^2 * 0 + \sum_{\{x: Prob(X=x)>0\}} g(x)^2 Prob(X = x) && (5)\\
&\Rightarrow 0 = \sum_{\{x: Prob(X=x)>0\}} g(x)^2 Prob(X = x) && (6)
\end{aligned}
$$

Suppose there exists $x' \in \{x : Prob(X = x) > 0\}$ such that $g(x') > 0$. Then, we would have:

$$\sum_{\{x: Prob(X=x)>0\}} g(x)^2 Prob(X = x) \geq g(x')^2 Prob(X = x') > 0$$

This is a contradiction of line 6, $0 = \sum_{\{x: Prob(X=x)>0\}} g(x)^2 Prob(X = x)$, so there cannot exist such an x'. Therefore, $g(x) = 0$ for all x such that $Prob(X = x) > 0$.

Part (a) (ii)

$$
\begin{aligned}
Prob(Y_1 = 1) &= \sum_x Prob(Y_1 = 1 | X = x) Prob(X = x) && (7)\\
&= \sum_x g(x) Prob(X = x) && (8)\\
&= \sum_{\{x: Prob(X=x)>0\}} g(x) Prob(X = x) && (9)\\
&= \sum_{\{x: Prob(X=x)>0\}} 0 * Prob(X = x) && (10)\\
&= 0 && (11)
\end{aligned}
$$

Line 13 uses the assumption $g(x) = 0$ for all x such that $Prob(X = x) > 0$. The same holds for Y_2 by symmetry.

Part (b) (i)

Solution:

Exchangeability follows because no matter how we permute Y_1 and Y_2, the probabilities remain the same:

$$0 = Prob(Y_2 = 0, Y_1 = 0) = Prob(Y_1 = 0, Y_2 = 0)$$
$$0 = Prob(Y_2 = 1, Y_1 = 1) = Prob(Y_1 = 1, Y_2 = 1)$$

$$1/2 = Prob(Y_2 = 0, Y_1 = 1) = Prob(Y_1 = 0, Y_2 = 1)$$
$$1/2 = Prob(Y_2 = 1, Y_1 = 0) = Prob(Y_1 = 1, Y_2 = 0)$$

Extra clarification:

We know that the events comprising all possible combinations of Y_1 and Y_2 sum to a total probability of 1:

$$\begin{aligned} 1 &= Prob(Y_1 = 1, Y_2 = 0) + Prob(Y_1 = 0, Y_2 = 1) + Prob(Y_1 = 0, Y_2 = 0) + Prob(Y_1 = 1, Y_2 = 1) \\ &= 1/2 + 1/2 + Prob(Y_1 = 0, Y_2 = 0) + Prob(Y_1 = 1, Y_2 = 1) \\ &= 1 + Prob(Y_1 = 0, Y_2 = 0) + Prob(Y_1 = 1, Y_2 = 1) \end{aligned}$$

This implies that
$$0 = Prob(Y_1 = 0, Y_2 = 0) = Prob(Y_1 = 1, Y_2 = 1)$$

Part (b) (ii)

Note from above that:
$$Prob(Y_1 = 1, Y_2 = 1) = 0$$

and
$$Prob(Y_1 = 1) = Prob(Y_1 = 1, Y_2 = 0) + Prob(Y_1 = 1, Y_2 = 1) = 1/2$$

Suppose for contradiction that there exists a discrete random variable X such that Y_1 and Y_2 are **conditionally iid** given X.

By part (a) (i), $Prob(Y_1 = 1, Y_2 = 1) = 0$ implies that $g(x) = 0$ for all x such that $Prob(X = x) > 0$.

By part (a) (ii), this further implies that $Prob(Y_1 = 1) = Prob(Y_2 = 1) = 0$.

This is a contradiction of the fact that $Prob(Y_1 = 1) = 1/2$.

So, there cannot exist any discrete random variable X such that Y_1 and Y_2 are **conditionally iid** given X.