STA 601/360 Homework 6

Yi Mi

17 十月, 2019

Exercise 1

Hoff 6.1

(a)

$$
\begin{aligned}
\theta &\sim gamma(a_\theta, b_\theta) \\
\gamma &\sim gamma(a_\gamma, b_\gamma) \\
E(\theta_A) &= E(\theta) \\
E(\theta_B) &= E(\theta * \gamma) = E(\theta)E(\gamma) \\
E(\theta_A\theta_B) &= E(\theta * \theta\gamma) = E(\theta^2\gamma) = E(\theta^2)E(\gamma) \\
E(\theta_A)E(\theta_B) &= E(\theta)E(\theta)E(\gamma) = [E(\theta)]^2E(\gamma) \\
Cov(\theta_A, \theta_B) &= E(\theta_A\theta_B) - E(\theta_A)E(\theta_B) = E(\theta^2)E(\gamma) - [E(\theta)]^2E(\gamma) \\
&= (E(\theta^2) - [E(\theta)]^2)E(\gamma) \\
&= Var(\theta)E(\gamma) \\
&= \frac{a_\theta}{b_\theta^2}E(\gamma) \\
&\neq 0 \\
\implies \theta_A &\text{ and } \theta_B \text{ are dependent.}
\end{aligned}
$$

In what situations is such a joint prior distribution justified?

(b)

$$
\begin{aligned}
p(\theta \mid y_A, y_B, \gamma) &\propto p(y_A, y_B \mid \theta, \gamma)p(\theta \mid \gamma) \\
&\propto p(y_A \mid \theta) * p(y_B \mid \theta, \gamma) * p(\theta) \\
&\propto \prod \theta^{y_{A,i}} e^{-\theta} * \prod (\theta\gamma)^{y_{B,i}} e^{-\theta\gamma} * \theta^{a_\theta - 1} e^{-b_\theta\theta} \\
&\propto \theta^{\Sigma y_{A,i}} e^{-n_A\theta} * \theta^{\Sigma y_{B,i}} e^{-n_B\theta\gamma} * \theta^{a_\theta - 1} e^{-b_\theta\theta} \\
&\propto \theta^{\Sigma y_{A,i} + \Sigma y_{B,i} + a_\theta - 1} * e^{-(n_A + n_B\gamma + b_\theta)\theta} \\
&\propto dgamma(\Sigma y_{A,i} + \Sigma y_{B,i} + a_\theta, n_A + n_B\gamma + b_\theta)
\end{aligned}
$$

(c)

$$
\begin{aligned}
p(\gamma \mid y_A, y_B, \theta) &\propto p(y_A, y_B \mid \theta, \gamma) p(\gamma \mid \theta) \\
&\propto p(y_B \mid \theta, \gamma) * p(\gamma) \\
&\propto \prod (\theta\gamma)^{y_{B,i}} e^{-\theta\gamma} * \gamma^{a_\gamma - 1} e^{-b_\gamma \gamma} \\
&\propto \gamma^{\Sigma y_{B,i}} e^{-n_B \theta \gamma} * \gamma^{a_\gamma - 1} e^{-b_\gamma \gamma} \\
&\propto \gamma^{\Sigma y_{B,i} + a_\gamma - 1} * e^{-(n_B \theta + b_\gamma)\gamma} \\
&\propto dgamma(\Sigma y_{B,i} + a_\gamma, n_B \theta + b_\gamma)
\end{aligned}
$$

(d)

```
a.theta = 2
b.theta = 1
ab.gamma = c(8,16,32,64,128)
#ab.gamma =8
yA=scan(file=url("http://www2.stat.duke.edu/~pdh10/FCBS/Exercises/mench
ild30bach.dat"))
yB=scan(file=url("http://www2.stat.duke.edu/~pdh10/FCBS/Exercises/mench
ild30nobach.dat"))

nA=length(yA)
nB=length(yB)

yA.sum=sum(yA)
yB.sum=sum(yB)

child.gibb = function(ab.gamma){

  a.gamma = b.gamma = ab.gamma
  S = 5000
  theta = mean(yA)
  gamma = mean(yB)/mean(yA)
  THETA = c()
  GAMMA = c()

  set.seed(20)
  for (s in 1:S) {

    theta = rgamma(1, a.theta + yA.sum + yB.sum, b.theta + nA +nB*gamma)
    gamma = rgamma(1, a.gamma + yB.sum, b.gamma + nB*theta)

    THETA[s] = theta
    GAMMA[s] = gamma
  }

  theta.A = THETA
  theta.B = THETA*GAMMA
  mean.theta.BA=mean(theta.B - theta.A)
```

```
}

child.mc = sapply(ab.gamma, child.gibb)
child.mc

## [1] 0.3868008 0.3389161 0.2748982 0.1970681 0.1309446
```

For prior distribution of γ, with a_γ and b_γ increasing, the posterior mean of $\theta_B - \theta_A$ decreases.

Exercise 2

Hoff 6.3

(a)

$$
\begin{aligned}
\beta &\sim N(0, \tau_\beta^2) \\
p(\beta \mid \mathbf{y}, \mathbf{x}, \mathbf{z}, c) &\propto p(\mathbf{z} \mid \mathbf{x}, \beta) * p(\beta) \\
&\propto \prod_i d\, norm(z_i, x_i\beta, 1) * dnorm(\beta, 0, \tau_\beta^2) \\
&= \prod_i \frac{1}{\sqrt{2\pi}} e^{-\frac{(z_i - x_i\beta)^2}{2}} * \frac{1}{\sqrt{2\pi\tau_\beta^2}} e^{-\frac{\beta^2}{2\tau_\beta^2}} \\
&\propto exp\{-\frac{1}{2}(\Sigma(z_i - x_i\beta)^2 + \frac{\beta^2}{\tau_\beta^2})\} \\
\text{Let } \frac{\beta^2}{\tau_\beta^2} + \Sigma z_i^2 & -2\beta\Sigma z_i x_i + \beta^2 \Sigma x_i^2 = a\beta^2 - 2b\beta + c, \\
\text{where } a &= \frac{1}{\tau_\beta^2} + \Sigma x_i^2,\ b = \Sigma z_i x_i,\ c = \Sigma z_i^2 \\
\Longrightarrow p(\beta \mid \mathbf{y}, \mathbf{x}, \mathbf{z}, c) &\propto exp\{-\frac{1}{2}(a\beta^2 - 2b\beta)\} \\
&\propto exp\{-\frac{1}{2}(\frac{\beta - b/a}{1/\sqrt{a}})^2\} \\
&\propto dnorm(\mu_n, \tau_n^2) \\
\text{where } \mu_n = \frac{b}{a} &= \frac{\Sigma z_i x_i}{1/\tau_\beta^2 + \Sigma x_i^2} \quad \text{and} \quad \tau_n^2 = \frac{1}{a} = \frac{1}{1/\tau_\beta^2 + \Sigma x_i^2}
\end{aligned}
$$

(b)

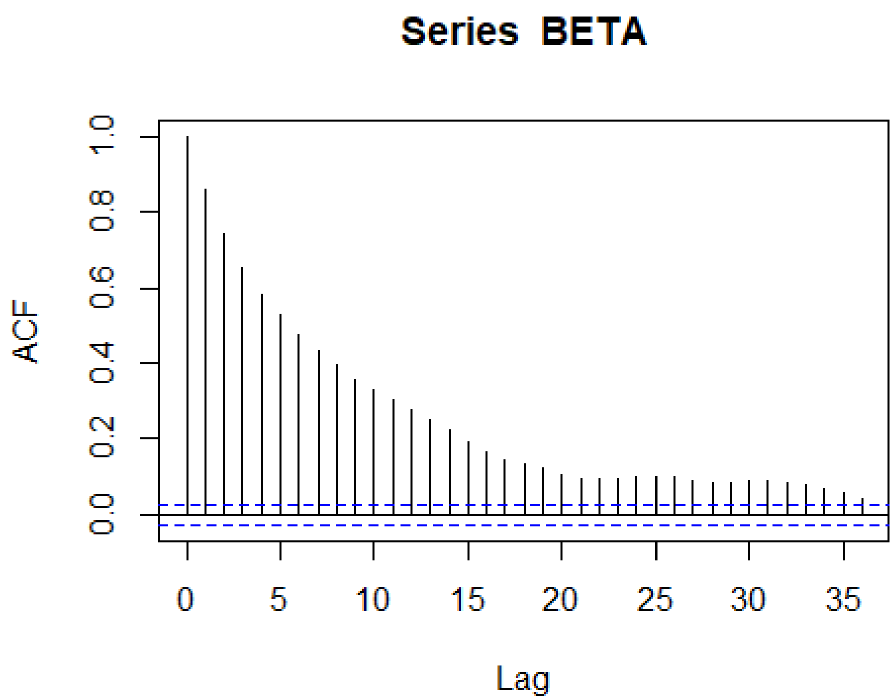
$$
\begin{aligned}
c &\sim N(0,\tau_c^2)\\
z_i &\sim N(x_i\beta,1)\\
p(c \mid \mathbf{y},\mathbf{x},\mathbf{z},\beta) &\propto p(c,\mathbf{y},\mathbf{x},\mathbf{z},\beta)\\
&\propto p(c) * p(\mathbf{z} \mid \mathbf{y},\mathbf{x},c,\beta)\\
p(z_i \mid y_i,x_i,c,\beta) &= \begin{cases} dnorm(z_i,x_i\beta,1)\mathbb{I}[z_i>c], & y_i=1\\ dnorm(z_i,x_i\beta,1)\mathbb{I}[z_i\le c], & y_i=0\end{cases}\\
\text{for all } y_i=1,\ \mathbb{I} &= \prod_i \mathbb{I}\,[z_i\colon y_i=1>c] = \mathbb{I}[min\{z_i\colon y_i=1\}>c]\\
\text{for all } y_i=0,\ \mathbb{I} &= \prod_i \mathbb{I}\,[z_i\colon y_i=0\le c] = \mathbb{I}[max\{z_i\colon y_i=0\}\le c]\\
\Longrightarrow p(\mathbf{z} \mid \mathbf{y},\mathbf{x},c,\beta) &= \prod_i d\,norm(z_i,x_i\beta,1) * \mathbb{I}[min\{z_i\colon y_i=1\}>c] * \mathbb{I}[max\{z_i\colon y_i=0\}\le c]\\
&= \prod_i d\,norm(z_i,x_i\beta,1) * \mathbb{I}[max\{z_i\colon y_i=0\}\le c<min\{z_i\colon y_i=1\}]\\
\Longrightarrow p(c \mid \mathbf{y},\mathbf{x},\mathbf{z},\beta) &\propto p(c) * \mathbb{I}[max\{z_i\colon y_i=0\}\le c<min\{z_i\colon y_i=1\}]\\
&\propto dnorm(c,0,\tau_c^2) * \mathbb{I}[max\{z_i\colon y_i=0\}\le c<min\{z_i\colon y_i=1\}]\\
\Longrightarrow p(c \mid \mathbf{y},\mathbf{x},\mathbf{z},\beta) &\quad \text{is a constrained normal density.}
\end{aligned}
$$

$$
\begin{aligned}
p(z_i \mid \mathbf{y},\mathbf{x},\mathbf{z}_{-i},\beta,c) &\propto p(z_i,\mathbf{y},\mathbf{x},\mathbf{z}_{-i},\beta,c)\\
&\propto p(z_i \mid y_i,x_i,\beta,c)\\
&= \begin{cases} dnorm(z_i,x_i\beta,1)\mathbb{I}[z_i>c], & y_i=1\\ dnorm(z_i,x_i\beta,1)\mathbb{I}[z_i\le c], & y_i=0\end{cases}
\end{aligned}
$$

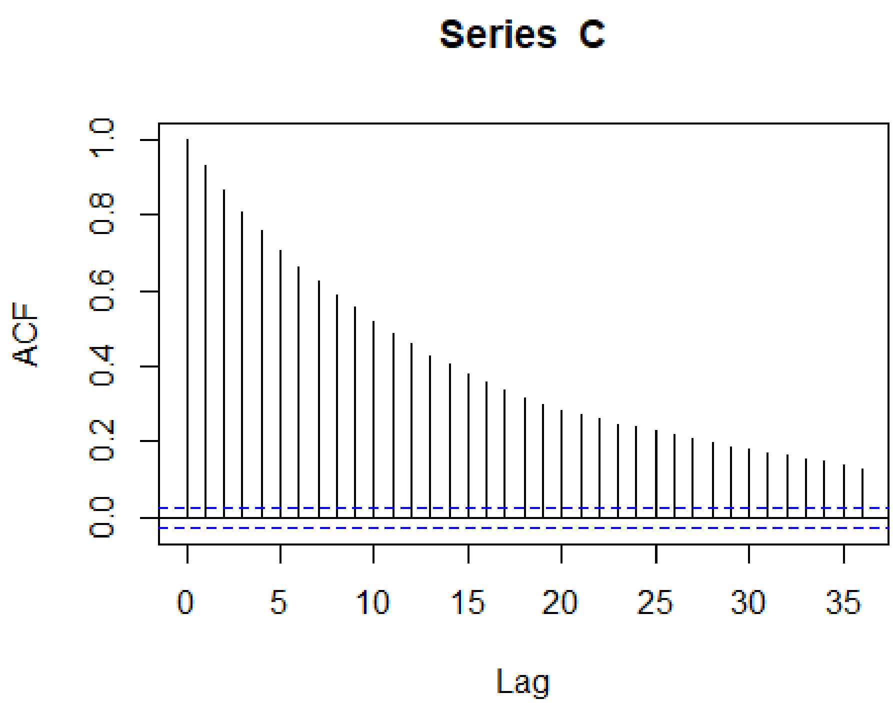
(c)

```
data = read.table(file=url("http://www2.stat.duke.edu/~pdh10/FCBS/Exerc
ises/divorce.dat"))

set.seed(1)

x = data[,1]
y = data[,2]
tau2.b=tau2.c=16
beta = rnorm(1, 0, sqrt(tau2.b))
c = rnorm(1, 0, sqrt(tau2.c))

#Gibbs
S=5000
Z=matrix(nrow = S,ncol=25)
BETA=c()
C=c()

for (s in 1:S) {
  # Z
```

```r
  z = c()
  for(i in 1:25){
    ez = beta*x[i]
    if(y[i] == 0){
      z[i] <- truncnorm::rtruncnorm(1, a = -Inf, b = c, mean = ez, sd =
 1)
      }
    if(y[i] == 1){
      z[i] <- truncnorm::rtruncnorm(1, a = c, b = Inf, mean = ez, sd =
1)
    }
  }

  # beta
  mun.b = sum(z*x)/(1/tau2.b + sum(x*x))
  taun.b = 1/(1/tau2.b + sum(x*x))
  beta = rnorm(1, mun.b, sqrt(taun.b))

  # c
  a.c = max(z[which(y==0)])
  b.c = min(z[which(y==1)])
  c = truncnorm::rtruncnorm(1, a = a.c, b = b.c, mean = 0, sd = sqrt(ta
u2.c))

  #store value
  Z[s,]=z
  BETA[s]=beta
  C[s]=c
}

acf(BETA)
```


Series BETA
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30
35
Lag

```r
acf(C)
```


Series C
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20
25
30
35
Lag

```r
# acf(Z)
```

```
effectiveSize(BETA)

##     var1
## 294.1021

effectiveSize(C)

##     var1
## 170.9775

effectiveSize(Z)

##      var1      var2      var3      var4      var5      var6      var
7
##  989.7319 2046.7853 4660.4046 1297.0140 3642.5214 4198.2453  630.158
4
##      var8      var9     var10     var11     var12     var13     var1
4
##  380.1552 3206.9486  790.4448 5000.0000  695.7843  475.5725  766.315
0
##     var15     var16     var17     var18     var19     var20     var2
1
##  704.4062 5000.0000  660.8594  709.0407 3196.3936 3182.1372  810.953
3
##     var22     var23     var24     var25
## 3261.8955  581.4448  357.9407 1569.6752
```

The autocorrelation plots shows that the autocorrelation is high, indicating that the mixing of the Markov chain is not good.

(d)

```
mean(BETA>0)

## [1] 0.999

quantile(BETA, c(0.025, 0.975))

##      2.5%     97.5%
## 0.1092148 0.6783859
```

Exercise 3

Hoff 7.3

(a)

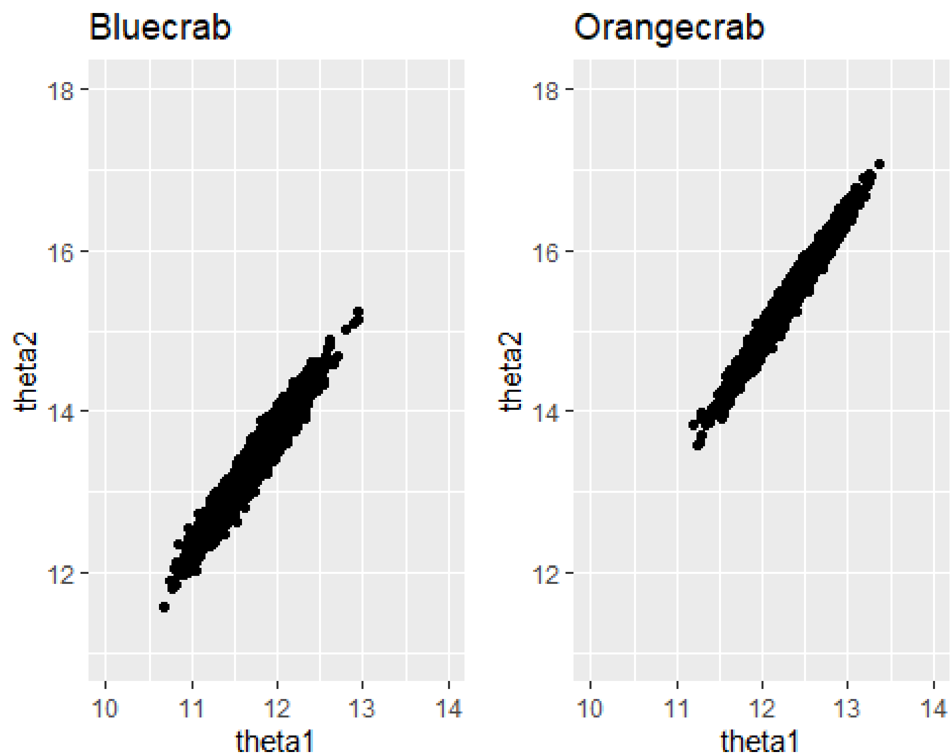
```
bluecrab = as.matrix(read.table(url('http://www2.stat.duke.edu/~pdh10/F
CBS/Exercises/bluecrab.dat')))
orangecrab = as.matrix(read.table(url('http://www2.stat.duke.edu/~pdh10
/FCBS/Exercises/orangecrab.dat')))
```

```
set.seed(1)

crab.mc = function(crab){

  n = dim(crab)[1]
  ybar = apply(crab, 2, mean)
  mu.0 = ybar
  S.0 = cov(crab)
  L.0 = sigma = S.0
  THETA = SIGMA = NULL

  v0 = 4
  S=5000

  #solve = Matrix::solve
  for(s in 1:S){
    L.n = Matrix::solve(Matrix::solve(L.0) + n*Matrix::solve(sigma))
    mu.n = L.n %*% (Matrix::solve(L.0) %*% mu.0 + n*Matrix::solve(sigma)
 %*% ybar)
    theta = rmvnorm(1, mu.n, L.n)

    S.theta = (t(crab) - c(theta)) %*% t(t(crab) - c(theta))
    S.n = S.0 + S.theta
    sigma = Matrix::solve(monomvn::rwish(v0 + n, Matrix::solve(S.n)))

    THETA = rbind(THETA, theta)
    SIGMA = rbind(SIGMA, c(sigma))
  }
  return(list(THETA, SIGMA))
}

bluecrab.mc = crab.mc(bluecrab)
orangecrab.mc = crab.mc(orangecrab)
```

(b)

```
theta.blue = data.frame(bluecrab.mc[[1]])
colnames(theta.blue) = c('theta1', 'theta2')
theta.orange = data.frame(orangecrab.mc[[1]])
colnames(theta.orange) = c('theta1', 'theta2')

gp.1 = ggplot(theta.blue, aes(x=theta1, y=theta2)) +
  geom_point() +
  labs(title = "Bluecrab", x = "theta1", y = "theta2") +
  scale_x_continuous(limits=c(10, 14)) +
  scale_y_continuous(limits=c(11, 18))
gp.2 = ggplot(theta.orange, aes(x=theta1, y=theta2)) +
  geom_point() +
  labs(title = "Orangecrab", x = "theta1", y = "theta2") +
  scale_x_continuous(limits=c(10, 14)) +
```

```
  scale_y_continuous(limits=c(11, 18))
gridExtra::grid.arrange(gp.1, gp.2, ncol=2)
```

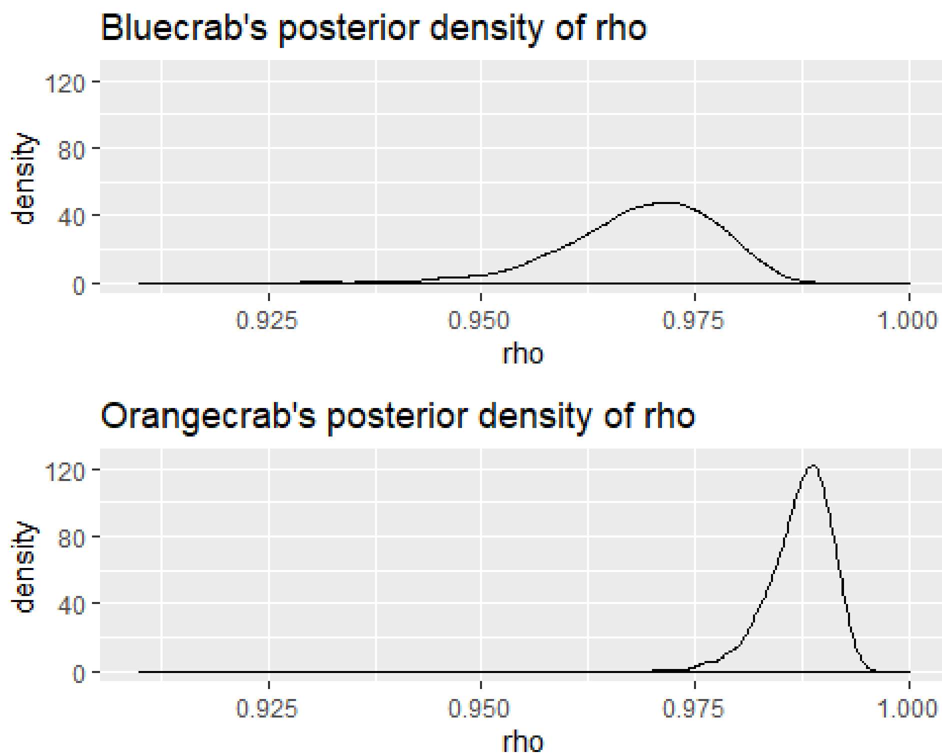


The scales of these two plots are the same. From the plot, we can see that orange crabs' both measurements of body depth and rear width are bigger than blue crabs'.

(c)

```
sigma.blue = data.frame(bluecrab.mc[[2]])
sigma.orange = data.frame(orangecrab.mc[[2]])

rho = function(sigma){
   sigma[2]/(sqrt(sigma[1]*sigma[4]))
}

cor.blue = apply(sigma.blue, MARGIN = 1, rho)
cor.orange = apply(sigma.orange, MARGIN = 1, rho)
df = data.frame(rbind(cbind(rho = cor.blue, crab = 'Bluecrab'),
                          cbind(rho = cor.orange, crab = 'Orangecrab')))

gp.3 = ggplot(data.frame(cor.blue), aes(x=cor.blue)) +
  geom_density() +
  labs(title = "Bluecrab's posterior density of rho", x = "rho", y = "d
ensity") +
  scale_x_continuous(limits=c(0.91, 1)) +
  scale_y_continuous(limits=c(0, 125))
gp.4 = ggplot(data.frame(cor.orange), aes(x=cor.orange)) +
```

```r
  geom_density() +
  labs(title = "Orangecrab's posterior density of rho", x = "rho", y =
"density") +
  scale_x_continuous(limits=c(0.91, 1)) +
  scale_y_continuous(limits=c(0, 125))
gridExtra::grid.arrange(gp.3, gp.4)
```




```r
mean(cor.blue<cor.orange)

## [1] 0.9854
```

The scales of these two plots are the same. From the plot and the MC approximation of $Pr(\rho_{\text{blue}} < \rho_{\text{orange}})$ is about 0.9854, we can see that the correlation of body depth and rear width of orange crabs tend to be more significant.