STA 601/360 Homework 5

Yi Mi

03 十月, 2019

Exercise 1

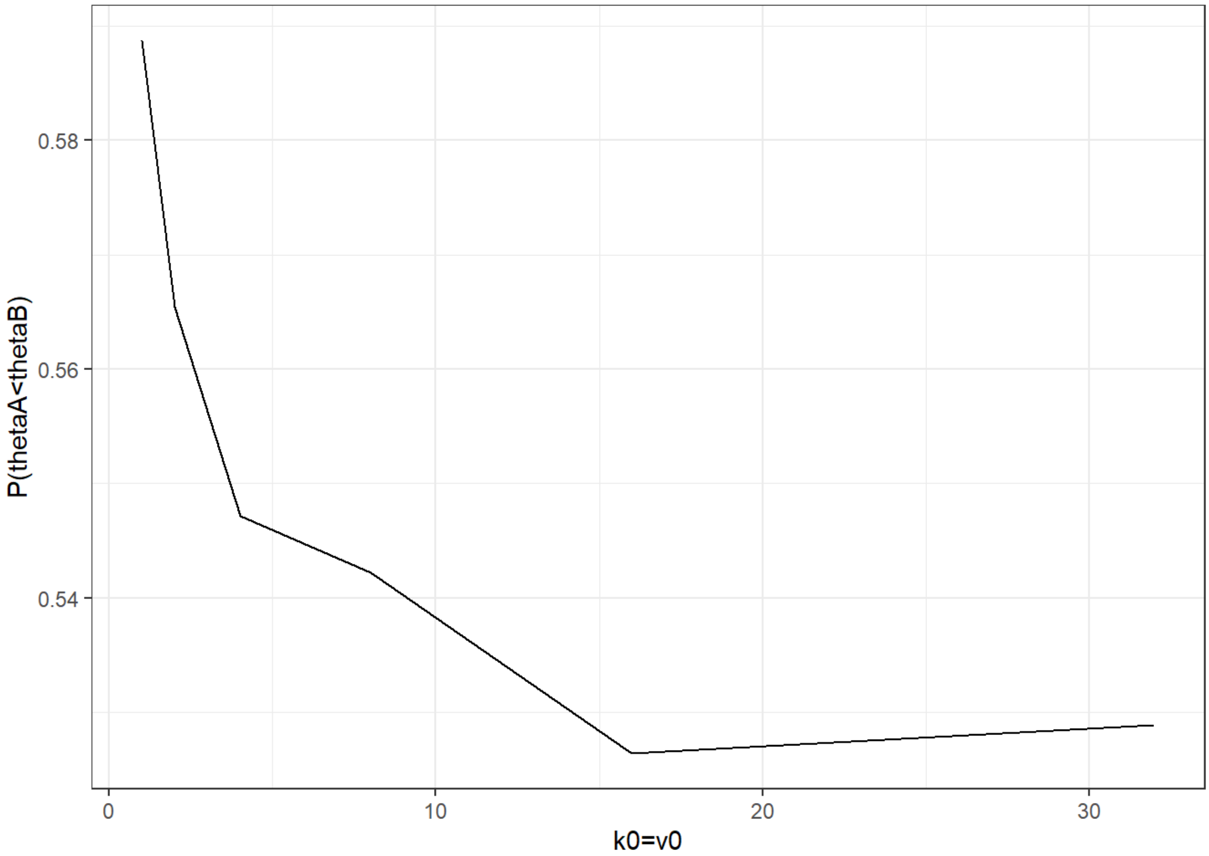
Hoff problem 5.2

Part (a)

Do this via Monte Carlo as the book recommends.

$$
\begin{aligned}
\frac{1}{\sigma^2} &\sim gamma(\frac{v_0}{2}, \frac{v_0\sigma_0^2}{2}) \\
\frac{1}{\sigma^2} \mid y_1, \ldots, y_n &\sim gamma(\frac{v_n}{2}, \frac{v_n\sigma_n^2}{2}) \\
v_n = v_0 + n,\ \sigma_n^2 &= 1/v_n * [v_0\sigma_0^2 + (n-1)s^2 + \frac{k_0 n}{k_n}(\bar{y} - \mu_0)^2] \\
\theta \mid y_1, \ldots y_n, \sigma^2 &\sim normal(\mu_n, \sigma^2/k_n) \\
k_n = k_0 + n,\ \mu_n &= \frac{k_0\mu_0 + n\bar{y}}{k_n}
\end{aligned}
$$

```
set.seed(20)
#
m.0=75
sigma2.0=100
yA=75.2
sA=7.3
yB=77.5
sB=8.1
nA=16
nB=16
#
theta <- function(v0, n, sigma2.0, s, k0, ybar, m.0){
  vn=v0+n
  kn=k0+n
  sigma2.n=(v0*sigma2.0^2+(n-1)*s^2+k0*n*(ybar-m.0)^2/kn)/vn
  m.n=(k0*m.0+n*ybar)/kn
  sigma2=1/rgamma(10000, vn/2, (vn*sigma2.n)/2)
  theta=rnorm(10000, m.n, sqrt(sigma2/kn))
}
#
mc1=c()
x=c(1,2,4,8,16,32)
#
for(i in 1:length(x)){
  theta.A=theta(x[i], nA, sigma2.0, sA, x[i], yA, m.0)
  theta.B=theta(x[i], nB, sigma2.0, sB, x[i], yB, m.0)
  mc1[i]=mean(theta.A<theta.B)
}
#
df1=data.frame(x=x,y=mc1)
ggplot(df1,aes(x=x,y=mc1))+geom_line()+labs(x='k0=v0', y='P(thetaA<thetaB)')
```



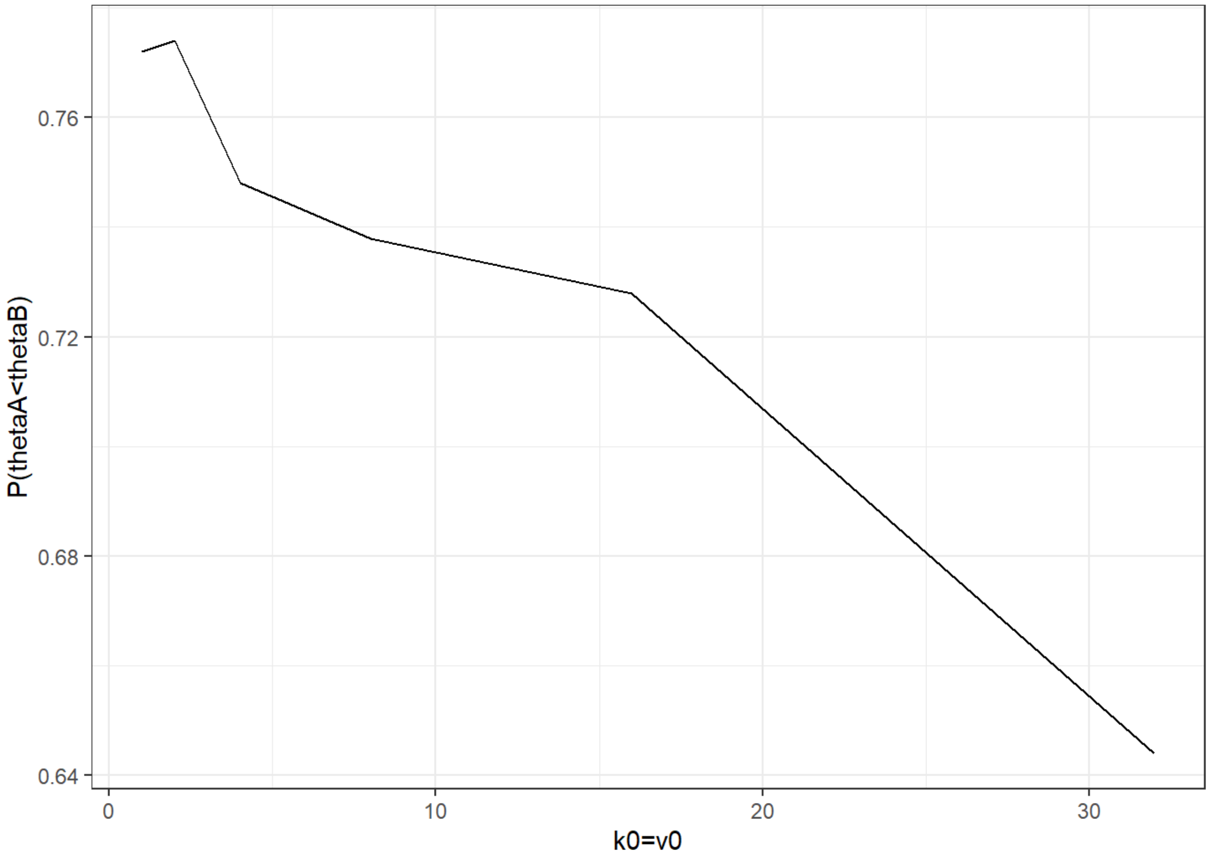
Part (b)

Repeat the above problem using the STAN code from lab. Comment on the quality of estimation when you run HMC for only a few hundred iterations. (You might need to edit the stan file to get the priors you want)

```
#
restandardize <- function(x, mean, sd){
  z <- (x - mean(x))/sd(x)
  return(sd*z + mean)
}
#
nA = nB = 16
ybarA = 75.2
sA = 7.3
ybarB = 77.5
sB = 8.1
#
yA = rnorm(nA) %>%
  restandardize(ybarA, sA)
yB = rnorm(nB) %>%
  restandardize(ybarB, sB)
#
x=c(1,2,4,8,16,32)
mc2=c()
# HMC
for(i in 1:length(x)){
  stan_res.A <- rstan::stan("hmc_norm.stan", data = list(y = yA, n = length(yA), m0=75, sigma0=100, v0
  =x[i], k0=x[i]),
                        chains = 1, iter = 600, warmup = 100, verbose = F, refresh = 0) %>%
            rstan::extract()
  stan_res.B <- rstan::stan("hmc_norm.stan", data = list(y = yB, n = length(yB), m0=75, sigma0=100, v0
  =x[i], k0=x[i]),
                        chains = 1, iter = 600, warmup = 100, verbose = F, refresh = 0) %>%
            rstan::extract()
  theta.A=stan_res.A$theta
  theta.B=stan_res.B$theta
  mc2[i]=mean(theta.A<theta.B)
}
```

```
## recompiling to avoid crashing R session
## recompiling to avoid crashing R session
## recompiling to avoid crashing R session
## recompiling to avoid crashing R session
## recompiling to avoid crashing R session
## recompiling to avoid crashing R session
## recompiling to avoid crashing R session
## recompiling to avoid crashing R session
## recompiling to avoid crashing R session
## recompiling to avoid crashing R session
## recompiling to avoid crashing R session
```

```
#
df2=data.frame(x=x,y=mc2)
ggplot(df2,aes(x=x,y=mc2))+geom_line()+labs(x='k0=v0', y='P(thetaA<thetaB)')
```



*The result of HMC estimation is very unstable for only a few hundred iterations.

Exercise 2

Jeffreys' prior for a normal (jointly for θ and σ^2) is proportional to $(\sigma^2)^{-\frac{3}{2}}$. Derive the posterior under this prior and state whether it is proper.

$$
\begin{aligned}
p(\theta,\sigma^2 \mid y_1,\dots,y_n) &= \frac{p(\theta,\sigma^2)p(y_1,\dots,y_n \mid \theta,\sigma^2)}{p(y_1,\dots,y_n)} \\
&\propto p(\theta,\sigma^2)p(y_1,\dots,y_n \mid \theta,\sigma^2) \\
&\propto (\sigma^2)^{-\frac{3}{2}}\sigma^{-n} exp(-\frac{\sum_{i=1}^n (y_i-\theta)^2}{2\sigma^2}) \\
&\propto (\sigma^2)^{-\frac{3+n}{2}} exp(-\frac{\sum_{i=1}^n (y_i-\theta)^2}{2\sigma^2}) \\
\Sigma(y_i-\theta)^2 &= \Sigma(\theta^2 - 2y\theta + y_i^2) \\
&= n\theta^2 - 2\theta\Sigma y_i + \Sigma y_i^2 \\
&= n(\theta^2 - 2\theta\frac{\Sigma y_i}{n}) + \Sigma y_i^2 \\
&= n(\theta^2 - 2\theta\frac{\Sigma y_i}{n} + (\frac{\Sigma y_i}{n})^2) - \frac{(\Sigma y_i)^2}{n} + \Sigma y_i^2 \\
&= n(\theta - \frac{\Sigma y_i}{n})^2 + (\Sigma y_i^2 - \frac{(\Sigma y_i)^2}{n}) \\
\implies p(\theta,\sigma^2 \mid y_1,\dots,y_n) &\propto (\sigma^2)^{-\frac{1}{2}}(\sigma^2)^{-(\frac{n}{2}+1)} exp(-\frac{n(\theta-\frac{\Sigma y_i}{n})^2 + (\Sigma y_i^2 - \frac{(\Sigma y_i)^2}{n})}{2\sigma^2}) \\
&\sim dNormal - inverse - gamma(\mu = \frac{\Sigma y_i}{n}, \lambda = n, \alpha = \frac{n}{2}, \beta = \frac{\Sigma y_i^2 - \frac{(\Sigma y_i)^2}{n})}{2}) \\
&\implies \text{The posterior distribution under this prior integrates to 1, indicating it is proper.}
\end{aligned}
$$

Exercise 3

3. Submit corrections for your Quiz 1 (the quiz is attached). (Please submit solutions to everything on the quiz, even if you got it correct. My recommendation would be to try to take the quiz again without outside material first and then go back to parts that you have missed). (You will receive up to half the points that you missed back on the quiz).

- The scan version of my Quiz 1 correction is attached.