HW4

suhy 2019010448

2023-04-03

5.10

a)

$$p(\tau|y) \propto p(\tau)V_\mu^{1/2}\prod_{j=1}^{J}(\sigma_j^2+\tau^2)^{-1/2}\exp\left(-\frac{(\bar{y}_{.j}-\hat{\mu})^2}{2(\sigma_j^2+\tau^2)}\right)$$

if $p(\tau) \propto \tau^{-1}$, when $\tau \to 0$, $p(\tau) \to \infty$ while $p(\tau|y)/p(\tau) \to C > 0$, so $\int_0^\infty p(\tau|y)d\tau \to \infty$, so the posterior density is improper.

b)

$$p(\tau|y) \propto V_\mu^{1/2}\prod_{j=1}^{J}(\sigma_j^2+\tau^2)^{-1/2}\exp\left(-\frac{(\bar{y}_{.j}-\hat{\mu})^2}{2(\sigma_j^2+\tau^2)}\right) \le V_\mu^{1/2}\prod_{j=1}^{J}(\sigma_j^2+\tau^2)^{-1/2} = \left(\sum_j\prod_{k\neq j}(\sigma_k^2+\tau^2)\right)^{-1/2} \le J^{1/2}\tau^{1-J}$$

when J>2, $p(\tau|y) \to 0$ when $\tau \to \infty$, and in a) we find that $p(\tau|y)/p(\tau) \to C > 0$, so $\int_0^\infty p(\tau|y)d\tau < \infty$, the posterior is proper.

5.12

$$E(\theta_j|\tau,y) = E[E[\theta_j|\mu,\tau,y]|\tau,y] = E\left[\frac{y_j/\sigma_j^2+\mu/\tau^2}{1/\sigma_j^2+1/\tau^2}|\tau,y\right] = \frac{y_j/\sigma_j^2+\hat{\mu}/\tau^2}{1/\sigma_j^2+1/\tau^2}, \hat{\mu} = E[\mu|\tau,y]$$

$$Var[\theta_j|\tau,y] = E[Var[\theta_j|\mu,\tau,y]|\tau,y]+Var[E[\theta_j|\mu,\tau,y]|\tau,y] = \frac{1}{1/\sigma_j^2+1/\tau^2}+\left(\frac{1/\tau^2}{1/\sigma_j^2+1/\tau^2}\right)^2V_\mu, V_\mu = Var[\mu|\tau,y]$$

5.13

a)

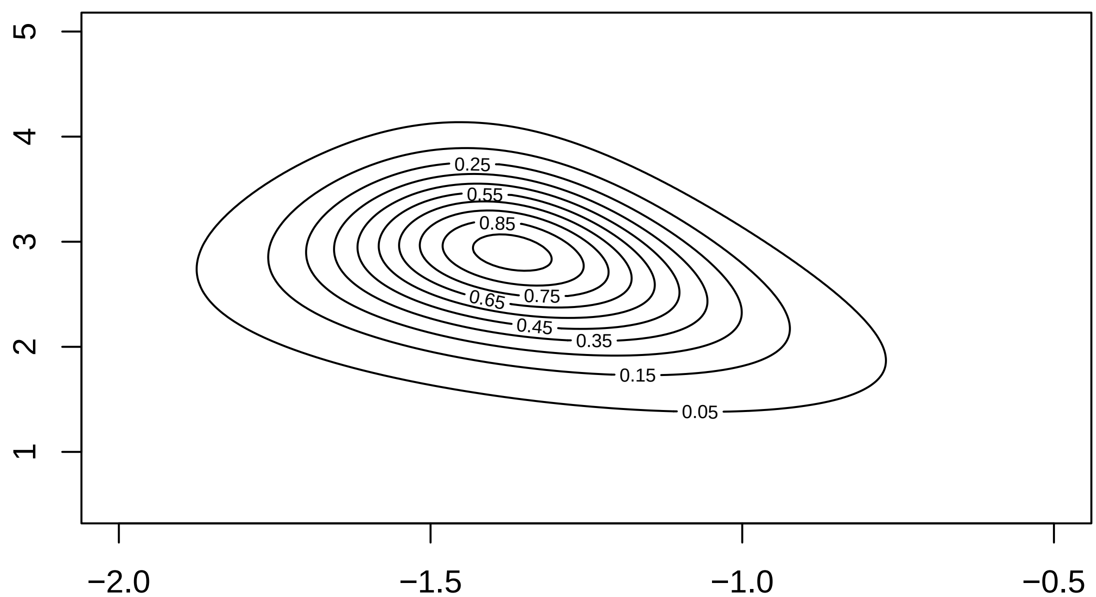
$$p(\theta,\alpha,\beta|y) = p(\alpha,\beta)\prod p(\theta_j|\alpha,\beta)p(y|\theta_j) = (\alpha+\beta)^{-5/2}\prod_{j=1}^{10}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta_j^{y_j+\alpha-1}(1-\theta_j)^{n_j-y_j+\beta-1}$$

b)

$$p(\alpha,\beta) = (\alpha+\beta)^{-5/2}\prod_{j=1}^{10}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma(\alpha+y_j)\Gamma(\beta+n_j-y_j)}{\Gamma(\alpha+\beta+n_j)}$$

```r
bike=c(16,9,10,13,19,20,18,17,35,55)
total=c(58,90,48,57,103,57,86,112,273,64) + bike
#evaluate the log posterior
logp <- function(u, v) { a <- (exp(u) * exp(v))/(exp(u) + 1);
b <- exp(v)/(exp(u) + 1);
```

```
J <- length(bike);
x <- J * (lgamma(a + b) - lgamma(a) - lgamma(b)) + log(a) + log(b) -
2.5 * log(a + b);
for(i in (1:J)) {
x <- x + lgamma(a + bike[i]) + lgamma(b + total[i] - bike[i]) -
lgamma(a + b + total[i]) }
x }
x <- c(-200:-50)/100
y <- c(50:500)/100
z <- outer(x,y,logp)
zz <- exp(z - max(z))
contour(x,y,zz,levels=c(0.05,0.15,0.25,0.35,0.45,0.55,0.65,0.75,0.85,0.95))
```



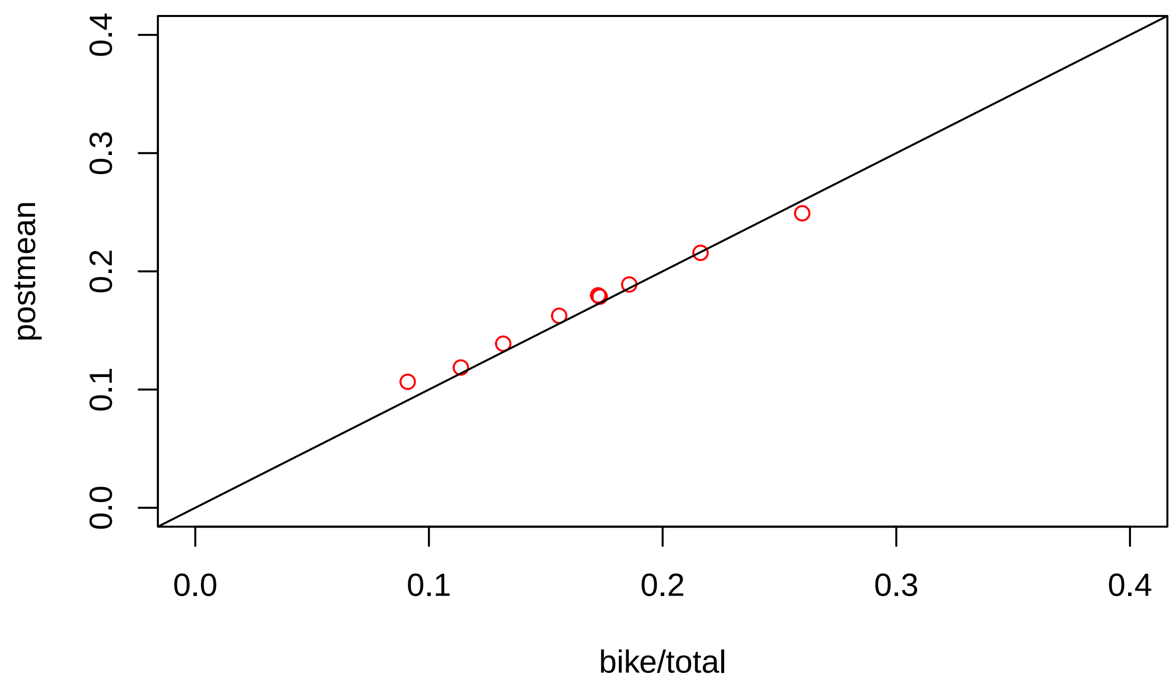
c)

```
m = 1000
zz <- zz/sum(zz)
p.x <- apply(zz,1,sum)
newx <- sample(x,m,replace=T,prob=p.x)
xid <- (newx+2.01)*100

newy <- rep(0,m)
for (i in (1:m)){
  newy[i] <- sample(y,1,prob=zz[xid[i],]/sum(zz[xid[i],]))}
```

```
newa <- (exp(newx)*exp(newy))/(exp(newx)+1)
newb <- exp(newy)/(exp(newx)+1)

J <- length(bike);
newth = matrix(0,nrow=m,ncol=J)
for (j in (1:J)) {
newth[,j] = rbeta(m,newa+bike[j],newb+total[j]-bike[j])
}
postmean <- colMeans(newth)
plot(bike/total,postmean,xlim=c(0,0.4),ylim=c(0,0.4),col='red')
abline(a=0,b=1)
```



raw proportion is close to the posterior distribution.

d)

```
theta.mean = rowMeans(newth)
(sort(theta.mean)[c(0.025*m,0.975*m)])
```

```
## [1] 0.1732404 0.2222978
```

e)

```
theta.new = rbeta(m,newa,newb)
(100*sort(theta.new)[c(0.025*m,0.975*m)])
```

```
## [1]  4.171863 50.742724
```

the interval is too wide for application.

6.6

a)

$$p_{pre}(y|\theta) \propto \theta^7(1-\theta)^{13}$$
$$p_{new}(y|\theta) = C_{12}^7\theta^7(1-\theta)^{13} \propto \theta^7(1-\theta)^{13}$$

two protocol are the have same $p(y|theta)$

b)

```
m=10000
T = rep(0,m)
for (i in 1:m){
theta = rbeta(1,8,14)
y = rbinom(1,1,theta)
while(length(which(y==0))<13){
y = c(y,rbinom(1,1,theta))
}
l = length(y)
T[i] = length(which(y[1:(l-1)]!=y[2:(l)]))
}
hist(T,,breaks=seq(-.5,max(T)+.5),cex=2)
abline(v=3,col='red')
```