Homework 2

Benjamin Sorenson

February 12, 2017

Chapter 2 Problems

- 5) For each of the following densities from Appendix A, provide a conjugate prior distribution for the unknown parameter(s0, eif one exists.
  a. $X \sim Bin(n, \theta)$, n known Given that $X \sim Bin(n, \theta)$,

  $$f(x|\theta) = \binom{n}{x}(1-\theta)^{n-x}\theta^x$$

  So, since $\binom{x}{k}$ does not depend on θ, we need to find a posterior $p(\theta|x)$ such that

  $$p(\theta|x) \propto \theta^{\alpha'}(1-\theta)^{\beta'} = f(x|\theta)\pi(\theta)$$

  For some α' and β'. This suggests that the conjugate prior $\pi(\theta|\eta)$ should be $Beta(\alpha, \beta)$. Which yields a posterior

  $$p(\theta|x) \propto (1-\theta)^{n-x}\theta^x(1-\theta)^{\beta-1}\theta^{\alpha-1} = (1-\theta)^{n-x+\beta-1}\theta^{x+\alpha-1}$$

  Setting $m(x) = B(x+\alpha, n-x+\beta)$, we see that

  $$p(\theta|x) \sim Beta(x+\alpha, n-x+\beta)$$

  b. $X \sim NegBin(r, \theta)$, r known
  c. $X \sim Mult(n, \boldsymbol{\theta})$, n, known
  d. $X \sim G(\alpha, \beta)$, α, known

- 7) Let θ be a univariate parameter of interest, and let $\gamma = g(\theta)$ be a 1-1 transform. Use (2.12) and (2.13) to show that (2.14) holds, i.e., that the Jeffreys prior is invariant under reperametrization. (*Hint:* What is the expectation of teh sol-called *score statistic* $\frac{d}{d\theta}\log f(\mathbf{x}|\theta)$)

- 9) Show that the Jeffreys prior based on teh binomial likelihood $f(x|\theta) = \binom{n}{x}\theta^x(1-\theta)^{n-x}$ is given by the $Beta(.5, .5)$ distribution

- 15) Suppose that $Y|\theta \sim G(1, \theta)$ (i.e., the *exponential* distribution with mean θ), and that $\theta \sim IG(1, \theta)$.
  a. Find the posterior distribution of θ.
  b. Find the posterior mean and variance of θ.
  c. Find the posterior mode of θ.
  d. Write down two integral equations that could be solved to find teh 95% equal-tail credible interval for θ.

Lab 3 Land Value Example

With NIG prior with parameters

$$\mu_\beta = (0,0,0,0)^T, V_\beta = 10^4(X^tX)^{-1}, a = b = 0.001$$

```r
# code to conduct posterior inference & prediction
# for the linear regression model for
# the land data using conjugate NIG prior

set.seed(810973206)

# first load "MASS" package which include the function
# "mvrnorm" to sample from multivariate normal dist.
library(MASS)

# read data from file
dir <- "~/datascience-masters/pubh7440/lab3/"
land.data <- read.table(file=file.path(dir, "land_data.txt"),header=T,sep="")
ls(land.data)
```

```
## [1] "X1" "X2" "X3" "Y"
```

```r
# define function to generate NITER samples of (beta,sigma^2)
# from the joint posterior using NIG prior with parameters "prior.para"
# and a given dataset "data"

post.sampling2 <- function(data, prior.para, NITER) {

    Y <- data[,'Y']
    X <- as.matrix(data[,2:4])
    X <- cbind(rep(1,times=length(Y)),X)
    n <- length(Y)
    p <- dim(X)[2]

    tXX <- t(X) %*% X
    tXX.inv <- solve(tXX)

    # extract the parameter in the NIG prior
    mu <- prior.para$mu
    V <- prior.para$V
    a <- prior.para$a
    b <- prior.para$b

    # calculate the posterior parameters
    V.star <- solve(solve(V) + tXX)
    mu.star <- V.star %*% (V %*% mu + t(X) %*% Y)
    a.star <- a + n/2
    b.star <- b + ( t(mu) %*% solve(V) %*% mu + t(Y) %*% Y
              - t(mu.star) %*% solve(V.star) %*% mu.star )/2


    # perform posterior sampling and return results
    sigma2 <- rep(NA, times = NITER)
    beta <- matrix(NA, nrow = NITER, ncol = p)
    colnames(beta) <- c('beta1','beta2','beta3','beta4')

    for (i in 1:NITER) {
        sigma2[i] <- 1/rgamma(1, a.star, rate=b.star)
        beta[i,] <- mvrnorm(1, mu.star, V.star)
```

```r
    }

    cbind(beta,sigma2)
}

# specify the prior parameters and collect posterior samples
X <- cbind(rep(1,times=nrow(land.data)),as.matrix(land.data[,2:4]))
prior.para = list(mu = rep(0,4),
                  V = 10000 * solve(t(X) %*% X),
                  a = 0.001,
                  b = 0.001
                  )

land.samples2 <- post.sampling2(land.data,prior.para,NITER=5000)


# define the function to compile summary statistics
sumstats <- function(vector){
    stats <- cbind(mean(vector),
                   sd(vector),
                   t(quantile(vector,c(.025,.5,.975))))
    names(stats) <- c('mean','sd','2.5%','50%','97.5%')
    stats
}

# summaries of the NEW samples given the NIG prior
t(apply(land.samples2,2,sumstats))
```

```
##              mean         sd       2.5%        50%      97.5%
## beta1   4.4473545 0.57205919  3.3566050  4.4527296  5.5860242
## beta2  -3.3638360 0.53962194 -4.4383020 -3.3657913 -2.2913978
## beta3   3.2291972 0.53078570  2.1837484  3.2252656  4.2388977
## beta4   1.5824647 0.60429297  0.4286227  1.5810607  2.7920131
## sigma2  0.5437668 0.03946117  0.4705079  0.5421748  0.6275217
```

```r
## Now we are to obtain the 95% posterior predictive interval

X.tilde <- c(1,0.870213,0.9453101,0.9198721)

post.predict <- function(X.tilde,post.samples) {
    NITER <- nrow(post.samples)
    y.tilde <- rep(NA,NITER)
    for (i in 1:NITER) {
        beta <- post.samples[i,1:4]
        sigma2 <- post.samples[i,5]
        y.mean <- X.tilde %*% beta
        y.tilde[i] <- rnorm(1,y.mean,sigma2)
    }
    y.tilde
}

# collect NEW predictive samples of y and calculate interval
pred.samples2 <- post.predict(X.tilde,land.samples2)
```

```r
exp(quantile(pred.samples2,c(.025,.975)))
```

```
##      2.5%     97.5%
##  140.1428 1210.9580
```

Bayes Factor bor the hypothesis:

$$H_0 : \beta_1 = 0$$

$$H_1 : \beta_1 \neq 0$$

```r
# R code to calculate Bayes Factor for hypothesis testing
# of H0: beta_1=0 vs H1: beta_1 <> 0
# for the land example

# load R package that contains the function 'dmvt'
# to calculate the pdf of MVSt
library(mvtnorm)

# read data from file
dir <- "~/datascience-masters/pubh7440/lab3/"
land.data <- read.table(file=file.path(dir,"land_data.txt"),header=T,sep="")
ls(land.data)
```

```
## [1] "X1" "X2" "X3" "Y"
```

```r
Y = land.data$Y

# Frist, specify likelihood and prior parameters of model M_0 and M_1
X.all = cbind(rep(1,times=nrow(land.data)),as.matrix(land.data[,2:4]))

M0.para <- list(X = X.all[,-4],
                mu = rep(0,3),
                V = 10^4 * solve(t(X.all[,-1]) %*% X.all[,-1]),
                a = 0.001,
                b = 0.001
                )


M1.para <- list(X = X.all,
                mu = rep(0,4),
                V = 10^4* solve(t(X.all) %*% X.all),
                a = 0.001,
                b = 0.001
                )

# define the function to calculate the posterior marginal
# of a given model m(y|M)
log.post.marg <- function(Y,M.para) {

    n <- length(Y)
    X <- M.para$X
    mu <- M.para$mu
    V <- M.para$V
```

```r
    a <- M.para$a
    b <- M.para$b

    y.nu <- 2 * a
    y.mean <- X %*% mu
    y.cov <- b / a * (diag(n) + X %*% V %*% t(X))

    dmvt(Y,y.mean,y.cov,df=y.nu)
}

# calculate the Bayes factor of M0 over M1
BF = exp(log.post.marg(Y,M0.para)-log.post.marg(Y,M1.para))
BF
```

```
## [1] 0.3161281
```

```r
1/BF
```

```
## [1] 3.163275
```

Given that the Bayes factor (`1/BF`) for the alternative hypothesis is 3.1632747, there is strong evidence to accept the alternative.