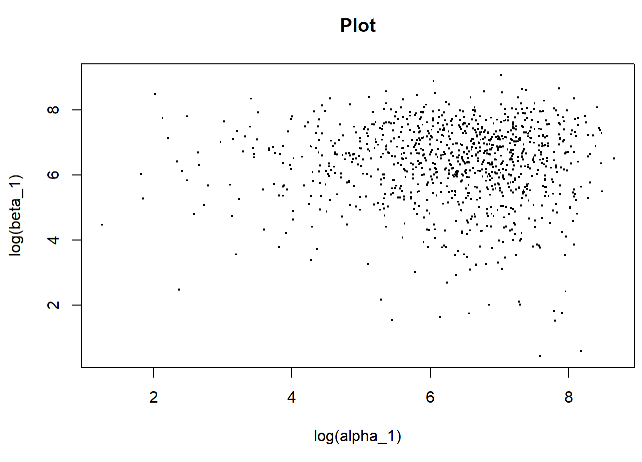
ADVANCED BAYESIAN MODELING - ASSIGNMENT 2

Shyam Shah

17/02/2021

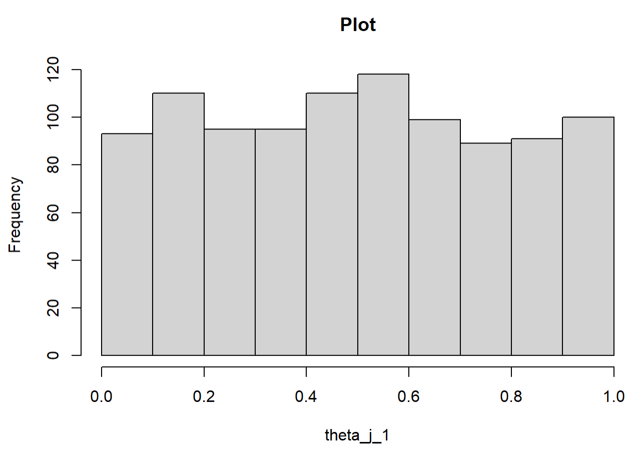
1 (a) (i)

```r
alpha_1 <- rexp(1000, rate=0.001)
beta_1 <- rexp(1000, rate=0.001)
plot (log(alpha_1),log(beta_1), main="Plot",pch=".", cex=2)
```



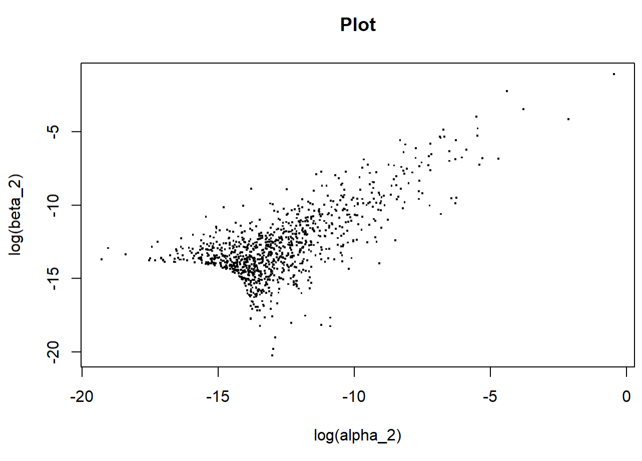
1 (a) (ii)

```r
theta_j_1 <- rbeta (1000, alpha_1, beta_1)
hist (theta_j_1, main="Plot")
```



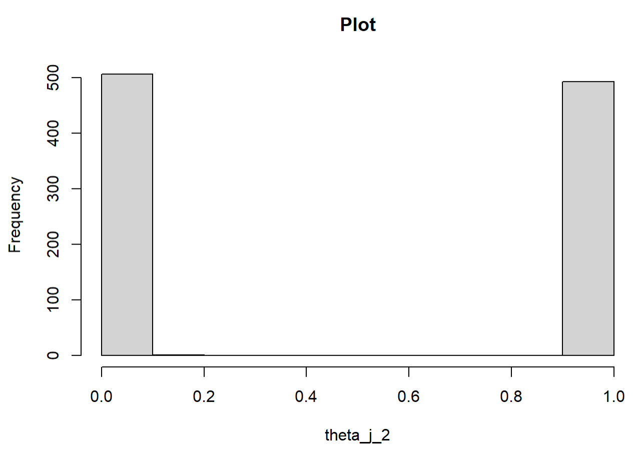
1 (b) (i)

```r
phi_1 <- runif(1000, 0, 1)
phi_2 <- runif(1000, 0, 1000)
alpha_2 <- phi_1/phi_2^2
beta_2 <- (1-phi_1)/phi_2^2
plot (log(alpha_2),log(beta_2), main="Plot",pch=".", cex=2)
```



1 (b) (ii)

```r
theta_j_2 <- rbeta (1000, alpha_2, beta_2)
hist (theta_j_2, main="Plot")
```



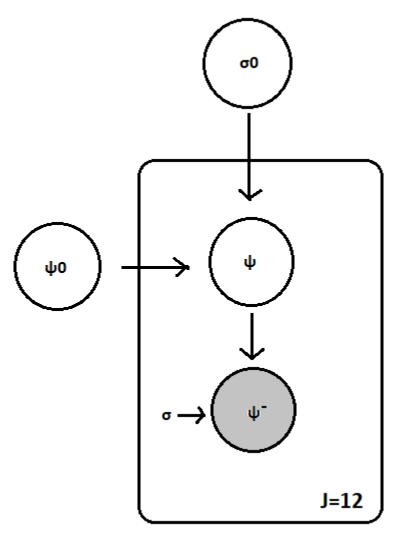
2 (a)

ψ_0 is meant to approximate an improper hyperprior that is meant to go from $-\infty$ to ∞. So the density it approximates is flat from $-\infty$ to ∞.

σ_0 is meant to approximate an improper hyperprior that is meant to go from 0 to ∞. So the density it approximates is flat from 0 to ∞.

These are the only hyperparameters.

2 (b)



All nodes have stochastic relationships with each other. I took $\bar{\psi}_j$ to be observed since it was computed from the observed data. σ_j is not in a circle because they are fixed and known. $\bar{\psi}_j$, σ_j and ψ_j are on the plate because they are vectors.

2 (c)

Below is what my JAGS model statement is.

```r
cat (readLines('./asgn2template.bug'), sep= '\n')
```

```
## model {
##   for (j in 1:length(sigma)) {
##     psihat[j] ~ dnorm(psi[j], 1/sigma[j]^2)
##     psi[j] ~ dnorm(psi0, 1/sigmasq0)
##   }
##
##   psi0 ~ dnorm(0, 1/1000^2)
##   sigma0 ~ dunif(0, 1000)
##
##   sigmasq0 <- sigma0^2
## }
```

2 (d)

```r
library(rjags) #import the rjag library
```

```
## Loading required package: coda
```

```
## Linked to JAGS 4.3.0
```

```
## Loaded modules: basemod,bugs
```

```r
data <- c(0.373, 0.116, 0.229, 0.117, 0.471, 0.120, 0.220, 0.239, 0.186, 0.328, 0.206, 0.254) #hard coded to ensu
re correctness
d <- list("sigma"=data) #create the list with the labels for jags to use
print(d)
```

```
## $sigma
##  [1] 0.373 0.116 0.229 0.117 0.471 0.120 0.220 0.239 0.186 0.328 0.206 0.254
```

```r
m <- jags.model("./asgn2template.bug", d, init=list(".RNG.name"="base::Wichmann-Hill", ".RNG.seed"=1)) #prepare t
he model with set seed
```

```
## Compiling model graph
##    Resolving undeclared variables
##    Allocating nodes
## Graph information:
##    Observed stochastic nodes: 0
##    Unobserved stochastic nodes: 26
##    Total graph size: 70
##
## Initializing model
```

```r
print (m)
```

```
## JAGS model:
##
## model {
##   for (j in 1:length(sigma)) {
##     psihat[j] ~ dnorm(psi[j], 1/sigma[j]^2)
##     psi[j] ~ dnorm(psi0, 1/sigmasq0)
##   }
##
##   psi0 ~ dnorm(0, 1/1000^2)
##   sigma0 ~ dunif(0, 1000)
##
##   sigmasq0 <- sigma0^2
## }
## Fully observed variables:
##  sigma
```

2 (e)

```r
update (m, 10000) #burn-in
x <- coda.samples(m, c("psi0", "sigmasq0"), n.iter=100000)
summary(x)
```

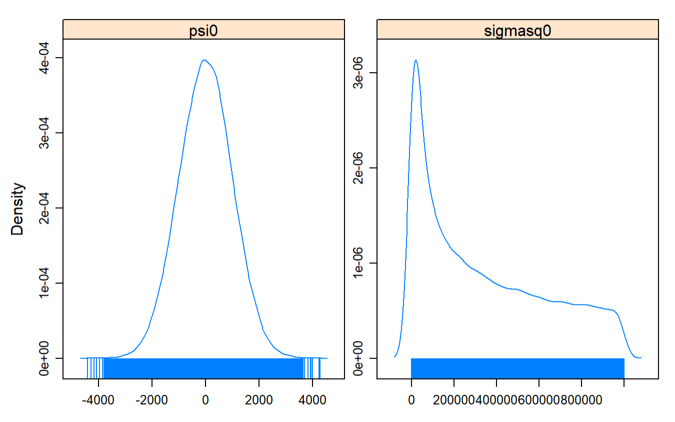
```
##
## Iterations = 10001:110000
## Thinning interval = 1
## Number of chains = 1
## Sample size per chain = 1e+05
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##               Mean       SD Naive SE Time-series SE
## psi0     3.696e+00    998.4    3.157          3.172
## sigmasq0 3.329e+05 298760.8  944.765        944.765
##
## 2. Quantiles for each variable:
##
##             2.5%   25%       50%      75%  97.5%
## psi0     -1949.0  -673 6.711e+00    678.4   1948
## sigmasq0   601.7 61650 2.489e+05 561158.5 952939
```

For ψ_0, the posterior expected value is 3.696, the posterior standard deviation is 998.4, and the 95% central posterior interval is (-1949,1948). For σ_0, the posterior expected value is 332900, the posterior standard deviation is 298760.8, and the 95% central posterior interval is (601.7,952939). The posterior densities are graphed below.

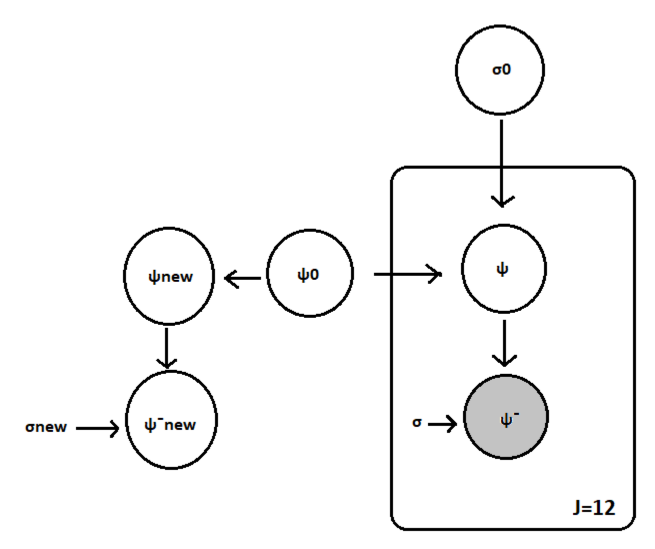
```r
require(lattice)
```

```
## Loading required package: lattice
```

```r
densityplot(x)
```



2 (f) (i)



I've added nodes for ψ_{new}, $\bar{\psi}_{new}$, and σ_{new}. $\bar{\psi}_{new}$ is not shaded because it is a new study so it is not observed, and they are not on a plate because it's a single study.

2 (f)(ii)

This is the modified JAGS model:

```r
cat (readLines('./asgn2templatepred.bug'), sep= '\n')
```

```
## model {
##   for (j in 1:length(sigma)) {
##     psihat[j] ~ dnorm(psi[j], 1/sigma[j]^2)
##     psi[j] ~ dnorm(psi0, 1/sigmasq0)
##   }
##
##   psi0 ~ dnorm(0, 1/1000^2)
##   sigma0 ~ dunif(0, 1000)
##
##   sigmasq0 <- sigma0^2
##
##   psihat_new ~ dnorm(psi_new, 1/sigma_new^2)
##   psi_new ~ dnorm(psi0, 1/sigmasq0)
##   ind <- psihat_new >= 2*sigma_new
## }
```

And these are the commands used to set it up:

```r
d <- list(sigma=data, sigma_new=c(0.2)) #add the known sigma for the new study
m <- jags.model("./asgn2templatepred.bug", d, init=list(".RNG.name"="base::Wichmann-Hill", ".RNG.seed"=1)) #prepa
re the model with set seed
```

```
## Compiling model graph
##    Resolving undeclared variables
##    Allocating nodes
## Graph information:
##    Observed stochastic nodes: 0
##    Unobserved stochastic nodes: 28
##    Total graph size: 77
##
## Initializing model
```

```r
update (m, 10000) #burn-in
x <- coda.samples(m, c("psihat_new","ind"), n.iter=100000)
```

#2 (f) (iii)

```r
summary(x)
```

```
##
## Iterations = 10001:110000
## Thinning interval = 1
## Number of chains = 1
## 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
## ind         0.4971    0.5 0.001581       0.001581
## psihat_new -3.3231 1157.3 3.659709       3.659709
##
## 2. Quantiles for each variable:
##
##             2.5%    25%    50%   75% 97.5%
## ind            0    0.0  0.000   1.0     1
## psihat_new -2271 -771.6 -7.206 765.8  2274
```

For $\bar{\psi}$, the posterior expected value is -3.3231, the posterior standard deviation is 1157.3, and the 95% central posterior interval is (-2271,2274).

2 (f) (iv)

The posterior predictive probability that the new estimated logs-odd ratio will be at least twice its standard error is 0.4971 from the table above.