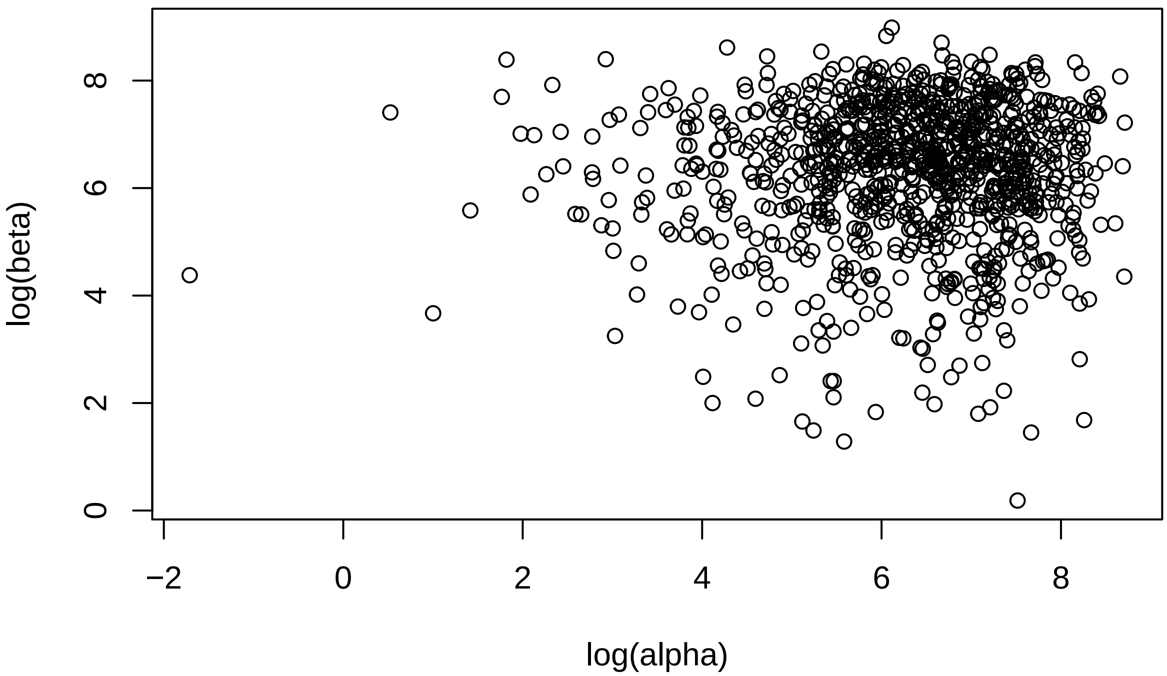
STAT 578 (Spring 2020) HW2 Solution

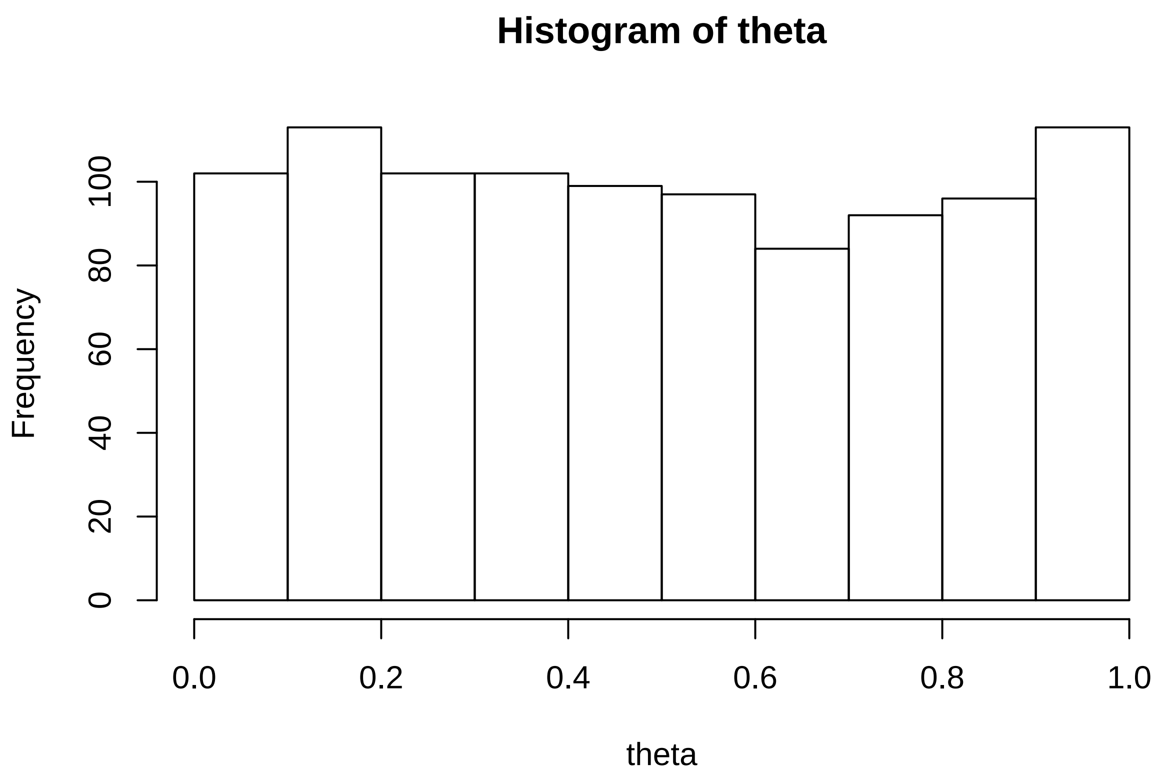
1. (a) (i)

```r
set.seed(578)
alpha <- rexp(1000, rate = 0.001)
beta  <- rexp(1000, rate = 0.001)
plot(log(alpha), log(beta))
```



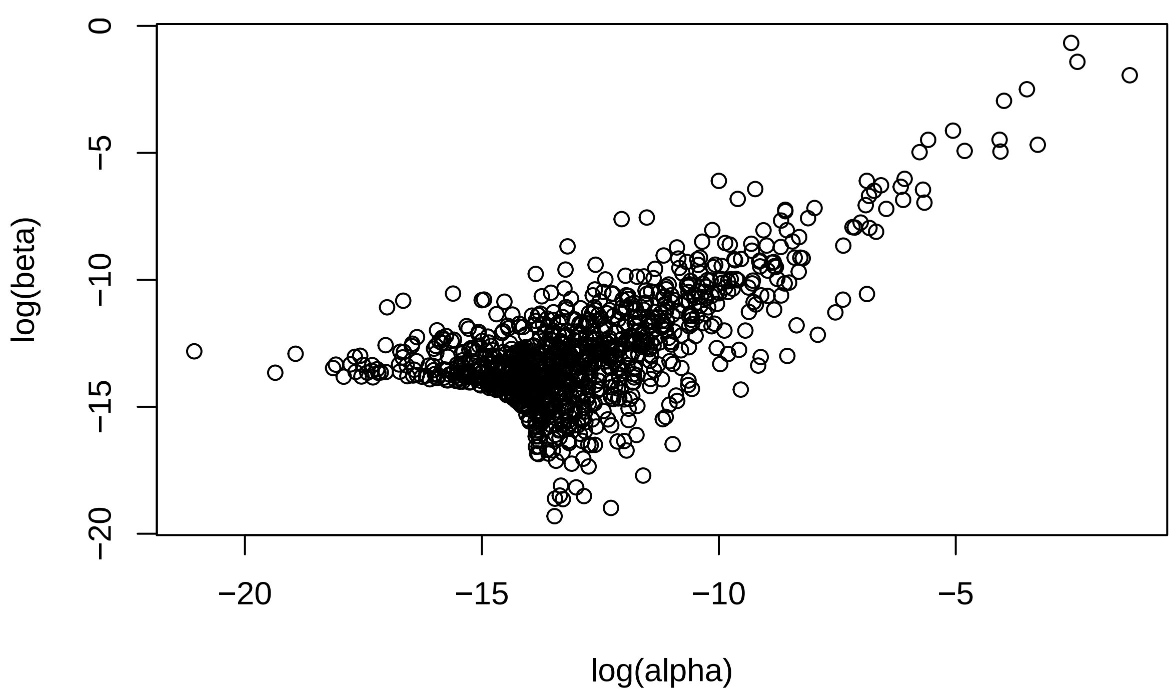
(ii)

```r
theta <- rbeta(1000, alpha, beta)
hist(theta)
```



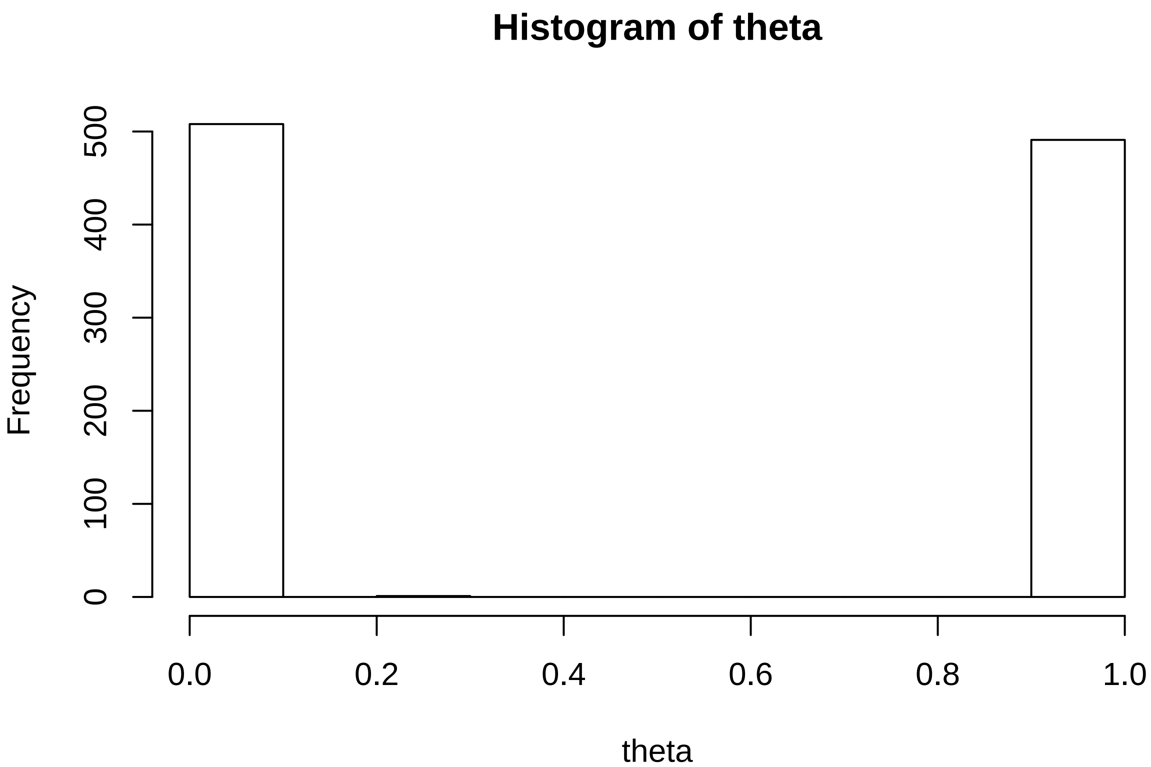
(b) (i)

```
set.seed(578)
phi1  <- runif(1000, 0, 1)
phi2  <- runif(1000, 0, 1000)
alpha <- phi1 / phi2^2
beta  <- (1-phi1) / phi2^2
plot(log(alpha), log(beta))
```



(ii)

```r
theta <- rbeta(1000, alpha, beta)
hist(theta)
```



2. (a) The improper priors being approximated are $p(\psi_0) \propto 1$ and $p(\sigma_0) \propto 1, \sigma_0 > 0$.

(b) See Figure 1.

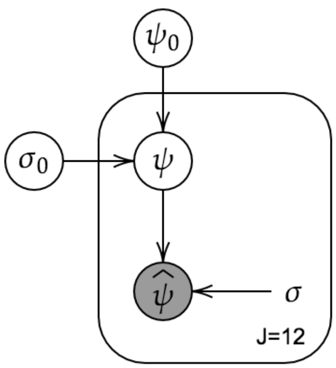


Figure 1: DAG of the Bayesian hierarchical model.

(c) JAGS code:

```
model {
    for (j in 1:12) {
        psihat[j] ~ dnorm(psi[j], 1.0/sigma[j]^2)
        psi[j] ~ dnorm(psi0, 1.0/sigmasq0)
    }
```

```
    psi0 ~ dnorm(0, 1.0/1000000)
    sigma0 ~ dunif(0, 1000)

    sigmasq0 <- sigma0^2
}
```

R code:

```
library(rjags)
tmp = scan(text=
"
1   1.055  0.373      5   1.068  0.471       9  0.507  0.186
2  -0.097  0.116      6  -0.025  0.120      10  0.000  0.328
3   0.626  0.229      7  -0.117  0.220      11  0.385  0.206
4   0.017  0.117      8  -0.381  0.239      12  0.405  0.254
")
d = as.data.frame(matrix(tmp, ncol=3, byrow=T)[, -1])
colnames(d) = c("psihat", "sigma")
inits <- list(.RNG.name="base::Super-Duper", .RNG.seed=578)
m <- jags.model("~/UIUC/STAT578_20Spring/HW2/hw2c.bug",
                d, inits)
```

(d) **Numerical summary:**

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

```
##
## Iterations = 11001:111000
## 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     0.2138 0.1330 0.0004205      0.0006442
## sigmasq0 0.1506 0.1309 0.0004141      0.0009869
##
## 2. Quantiles for each variable:
##
##              2.5%     25%    50%    75%  97.5%
## psi0     -0.03238 0.12812 0.2071 0.2934 0.4947
## sigmasq0  0.01880 0.06817 0.1161 0.1906 0.4869
```

- ψ_0:

posterior mean: **0.2138**

posterior sd: **0.1330**

95% central posterior intervals: **(-0.03238, 0.4947)**

- σ_0^2:

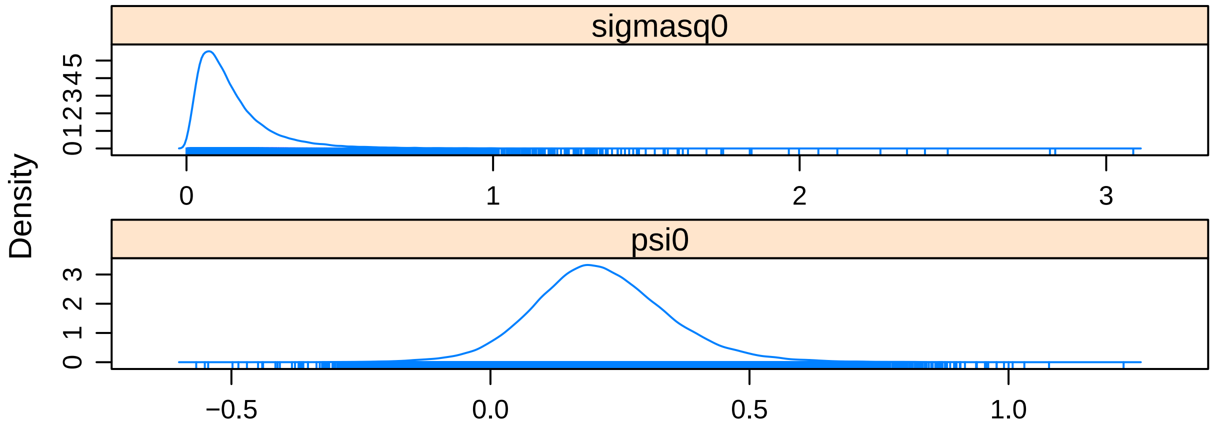
posterior mean: **0.1506**

posterior sd: **0.1309**

95% central posterior intervals: **(0.01880, 0.4869)**

Posterior densities:

```
library(lattice)
densityplot(x)
```



(e) (i) See Figure 2.

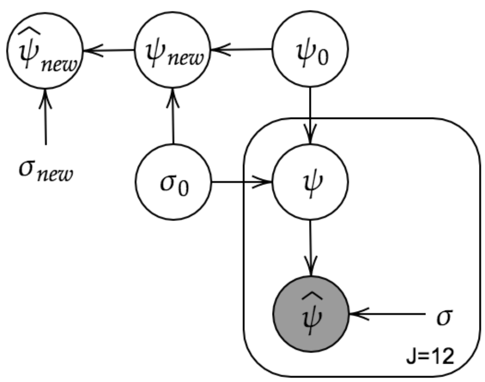


Figure 2: DAG of with new nodes.

(ii) JAGS code:

```
model {
    for (j in 1:12) {
        psihat[j] ~ dnorm(psi[j], 1.0/sigma[j]^2)
        psi[j] ~ dnorm(psi0, 1.0/sigmasq0)
    }

    psi0 ~ dnorm(0, 1.0/1000000)
    sigma0 ~ dunif(0, 1000)

    sigmasq0 <- sigma0^2

    psihat.new ~ dnorm(psi.new, 1/sigma.new^2)
    psi.new ~ dnorm(psi0, 1.0/sigmasq0)
```

```
    indicator <- psihat.new > 2*sigma.new
}
```

R code:

```
m2 <- jags.model("~/UIUC/STAT578_20Spring/HW2/hw2e.bug",
                 c(as.list(d), sigma.new=0.125),
                 inits)
```

(iii)

```
update(m2, 10000)
x2 <- coda.samples(m2, c('psihat.new', 'indicator'), n.iter=100000)
summary(x2)
```

```
##
## Iterations = 11001:111000
## 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
## indicator  0.4490 0.4974 0.001573       0.001705
## psihat.new 0.2126 0.4258 0.001346       0.001430
##
## 2. Quantiles for each variable:
##
##               2.5%      25%    50%    75% 97.5%
## indicator   0.0000  0.00000 0.0000 1.0000 1.000
## psihat.new -0.6298 -0.04151 0.2039 0.4608 1.098
```

- the new $\hat{\psi}$:

posterior mean: **0.2126**

posterior sd: **0.4258**

95% central posterior predictive intervals: **(-0.6298, 1.098)**

(iv) As shown in the output from (iii), the estimated posterior predictive probability is **0.4490**.