STAT 578 - Fall 2019 - Assignment 2

Frederick (Eric) Ellwanger - fre2

September 28, 2019

Excersise 1

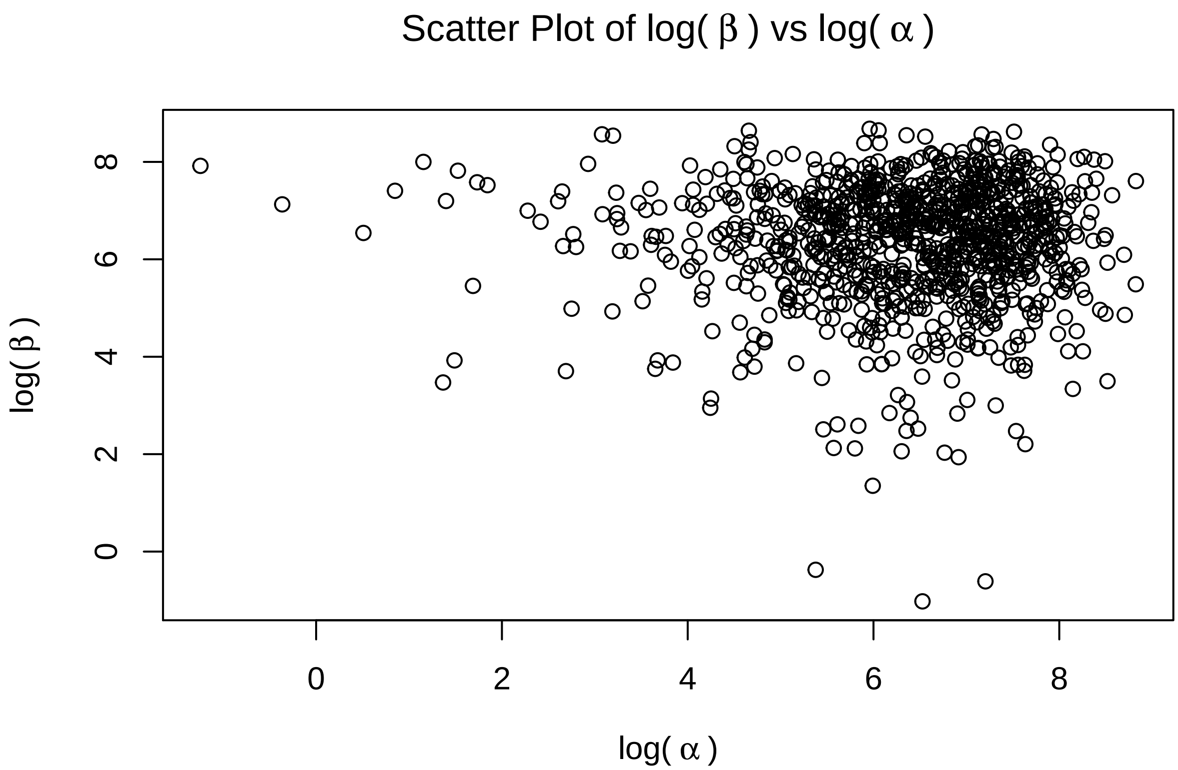
1. Consider the two different hyperprior formulations for the binomial hierarchical model of Lesson 3.2: Hierarchical Modeling Fundamentals. This exercise shows how different those priors are.

(1)(a) The first prior formulation was:

$$\theta_j|\alpha, \beta \sim Beta(\alpha, \beta)$$
$$\alpha, \beta \sim iid \;\; Expon(0.001)$$

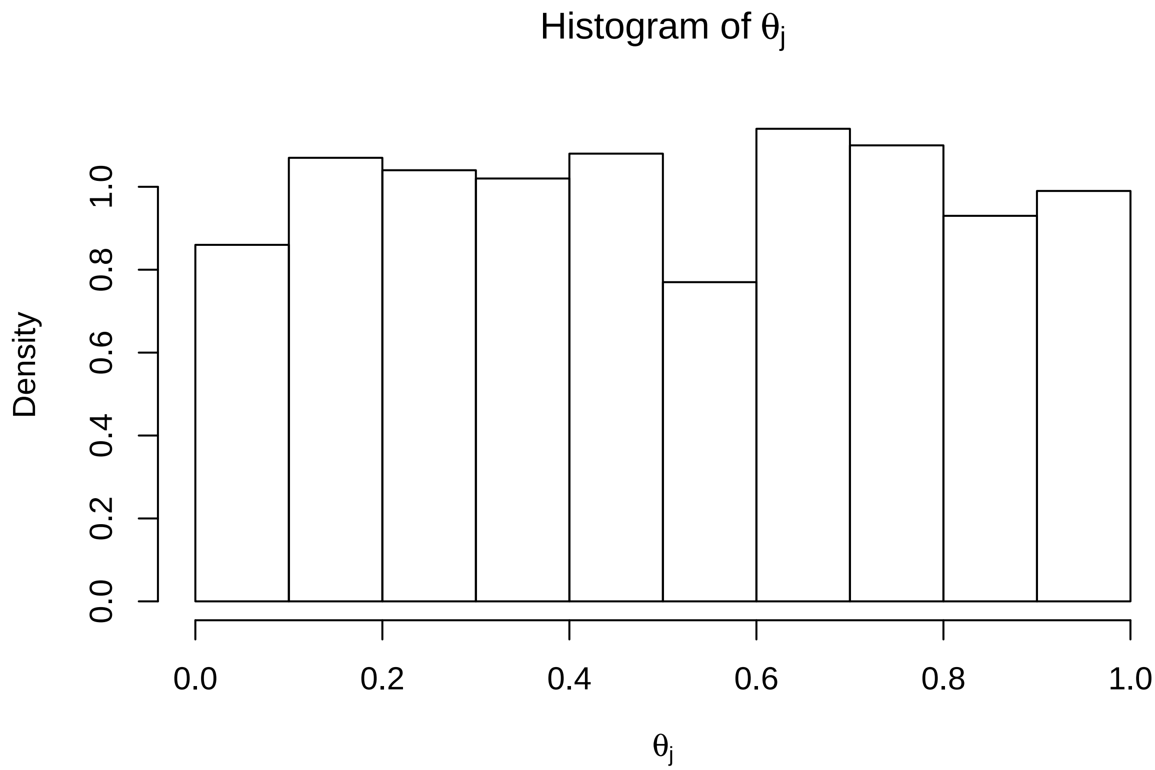
(1)(a)(i) Independently simulate 1000 pairs (α, β) from their hyperprior, and produce a scatterplot of $\log(\beta)$ versus $\log(\alpha)$.

```r
set.seed(19690223)

#Simulate 1000 draws of alpha and beta
alpha = rexp(1000, rate = 0.001)
beta = rexp(1000, rate = 0.001)

plot(log(alpha), log(beta), ylab = bquote("log("~beta~")"), xlab = bquote("log("~alpha~")"),
                            main = bquote("Scatter Plot of log("~beta~") vs log("~alpha~")"))
```



(1)(a)(ii) Using the simulated pairs (α, β), forward-simulate θ_j , and produce a histogram of the result (an approximation of its marginal prior).

```
thetaj = rbeta(1000, alpha, beta)
hist(thetaj, freq = FALSE, xlab = bquote(theta[j]), main = bquote("Histogram of"~theta[j]))
```

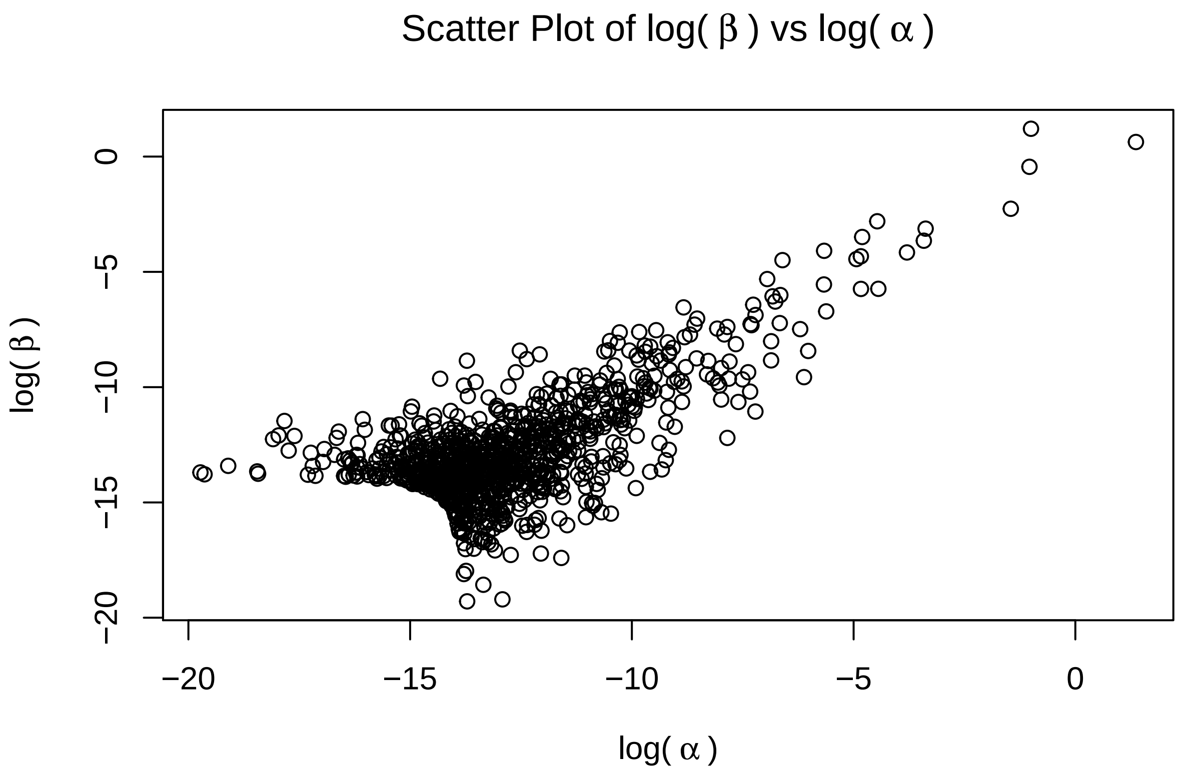


(1)(b) The second prior formulation was

$$\begin{aligned} \theta_j | \alpha, \beta &\sim Beta(\alpha, \beta) \\ \alpha = \frac{\phi_1}{\phi_2^2} \qquad \beta &= \frac{(1-\phi_1)}{\phi_2^2} \\ \phi_1 \sim U(0,1) \qquad \phi_2 &\sim U(0, 1000) \end{aligned} \tag{1}$$

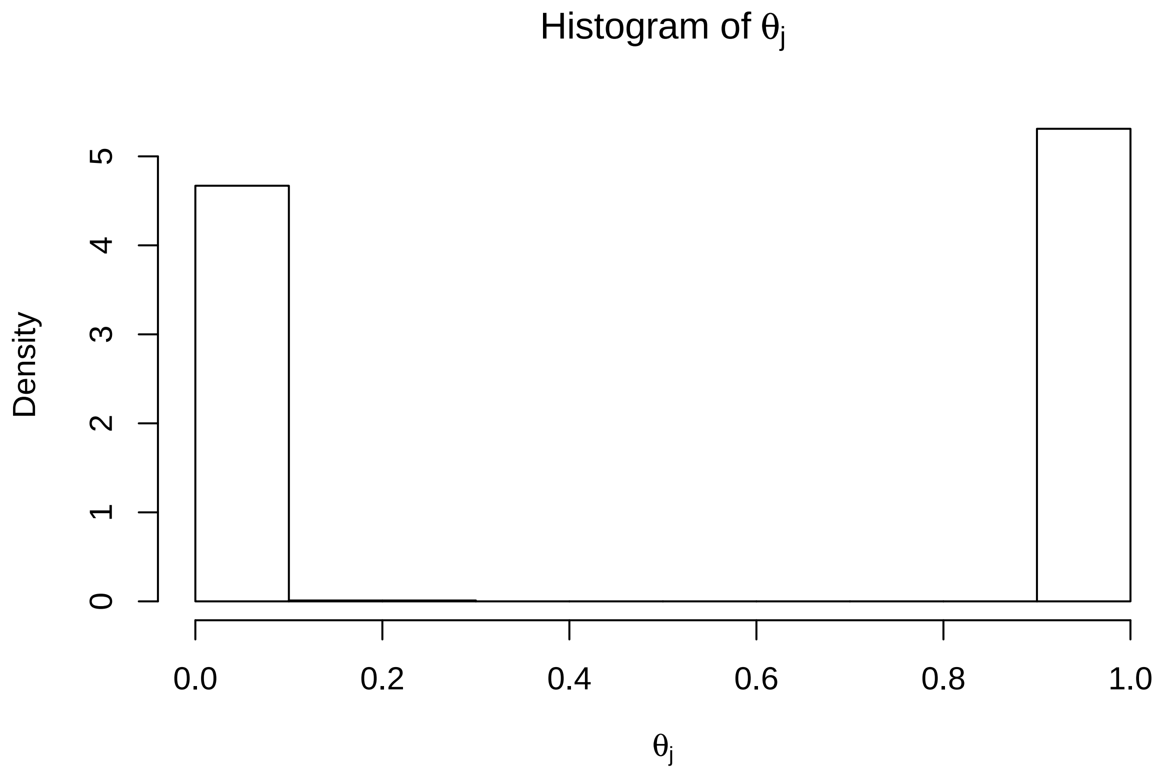
(1)(b)(i) Independently simulate 1000 pairs (α, β) from their hyperprior, and produce a scatterplot of $\log(\beta)$ versus $\log(\alpha)$.

```
set.seed(19690223)

#Simulate 1000 draws of phi1 and phi2
phi1 = runif(1000, 0, 1)
phi2 = runif(1000, 0, 1000)

alpha = phi1/phi2^2
beta = (1 - phi1)/phi2^2

plot(log(alpha), log(beta), ylab = bquote("log("~beta~")"), xlab = bquote("log("~alpha~")"),
                            main = bquote("Scatter Plot of log("~beta~") vs log("~alpha~")"))
```



(1)(b)(ii) Using the simulated pairs (α, β), forward-simulate θ_j , and produce a histogram of the result (an approximation of its marginal prior).

```
thetaj = rbeta(1000, alpha, beta)
hist(thetaj, freq = FALSE, xlab = bquote(theta[j]), main = bquote("Histogram of"~theta[j]))
```



Excersise 2

(2) Consider this Bayesian hierarchical model:

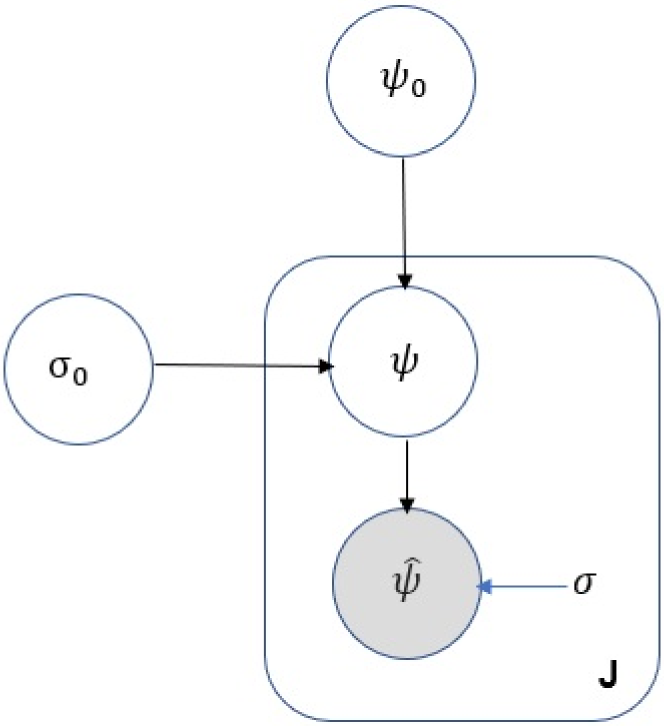
$$\begin{aligned}
\hat{\psi}_j|\psi_j \sim indep.\mathcal{N}(\psi_j, \sigma_j^2) \qquad j = 1, ..., 12 \\
\psi_j|\psi_0, \sigma_0 \sim iid \quad \mathcal{N}(\psi_0, \sigma_0^2) \qquad j = 1, ..., 12 \\
\psi_0 \sim \ \mathcal{N}(0, 1000^2) \\
\sigma_0 \sim U(0, 1000)
\end{aligned} \tag{2}$$

(2)(a) Specify improper densities that the proper hyperpriors given above appear to be approximating. (Which parameters are the hyperparameters?)

The hyperparameters are: ψ_0 **and** σ_0
The improper densities of the hyperpriors appear to be approximating:

$$\begin{aligned}
p(\psi_0) \propto 1 \qquad (flat \ \ on \ \ -\infty < \ \ \psi_0 \ \ < \infty) \\
p(\sigma_0) \propto 1 \qquad (flat \ \ on \ \ 0 < \ \ \sigma_0 \ \ < \infty)
\end{aligned} \tag{3}$$

(2)(b) Draw a directed acyclic graph (DAG) appropriate for this model. (Use the notation introduced in lecture, including plates.) You may draw it neatly by hand or use software.



(2)(c) Using the template asgn2template.bug provided on the course website, form a JAGS model statement (corresponding to your graph). Also, set up any R (rjags) statements appropriate for creating a JAGS model. [Remember: JAGS dnorm uses precisions, not variances!]

```
model {
  for (j in 1:length(psi_hat)) {
    psi_hat[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
}
```

```
library(rjags)
d = read.table("asgn2data.txt", header=TRUE)
m1 = jags.model("asgn2template.bug", d)
```

```
## Compiling model graph
##    Resolving undeclared variables
##    Allocating nodes
## Graph information:
##    Observed stochastic nodes: 12
```

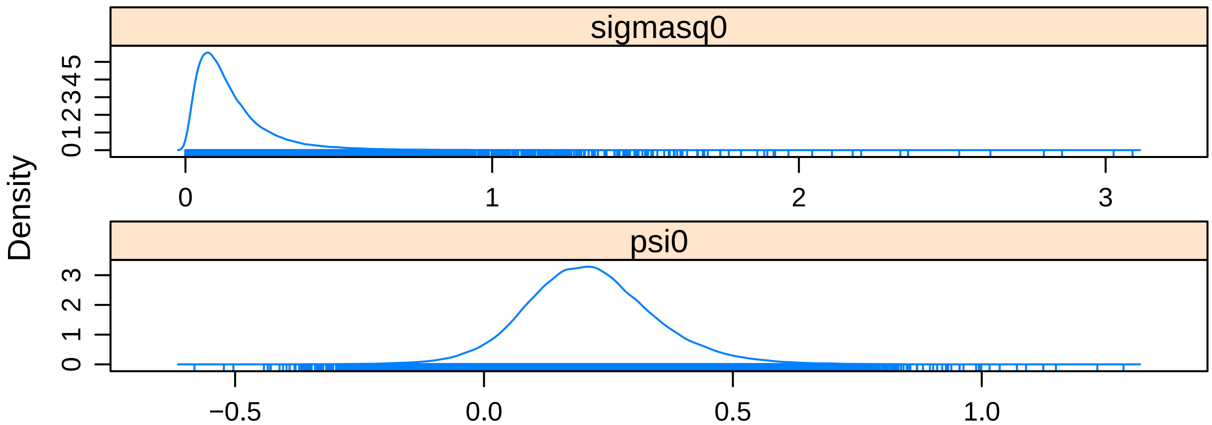
```
##    Unobserved stochastic nodes: 14
##    Total graph size: 70
##
## Initializing model
```

(2)(d) Run at least 10,000 iterations of burn-in, then 100,000 iterations to use for inference. For both ψ_0 and σ_0^2 (not σ_0), produce a posterior numerical summary and also graphical estimates of the posterior densities. Explicitly give the approximations of the posterior expected values, posterior standard deviations, and 95% central posterior intervals. (Just showing R output is not enough!)

```r
update(m1, 10000)   #Burn-in

x1 = coda.samples(m1, c("psi0", "sigmasq0"), n.iter = 100000)
summary(x1)
```

```
##
## Iterations = 11001:111000
## Thinning interval = 1
## Number of chains = 1
## Sample size per chain = 100000
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##           Mean    SD Naive SE Time-series SE
## psi0     0.214 0.133 0.000421       0.000639
## sigmasq0 0.151 0.135 0.000427       0.000988
##
## 2. Quantiles for each variable:
##
##             2.5%    25%   50%   75% 97.5%
## psi0     -0.0337 0.1273 0.207 0.293 0.496
## sigmasq0  0.0185 0.0678 0.116 0.191 0.497
```

```r
require(lattice)
densityplot(x1[,c("psi0", "sigmasq0")])
```



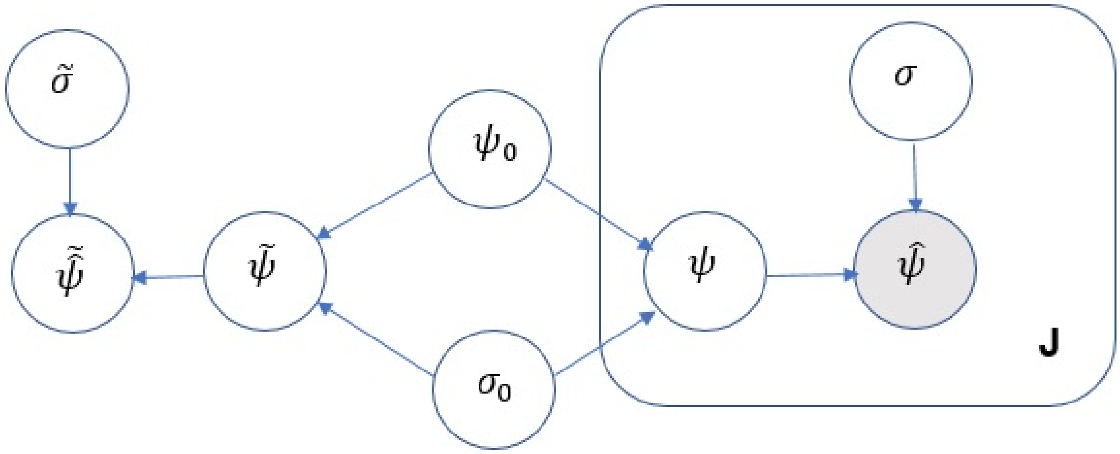
posterior expected value of $\psi_0 \approx$ 0.214
posterior expected value of $\sigma_0^2 \approx$ 0.151

posterior standard deviation of $\psi_0 \approx$ 0.133
posterior standard deviation of $\sigma_0^2 \approx$ 0.135

95% central posterior interval for $\psi_0 \approx$ (-0.034, 0.496)
95% central posterior interval for $\sigma_0^2 \approx$ (0.018, 0.497)

(2)(e) Suppose a new case-control study is to be performed, and assume that its log-odds standard error (new σ) will be 0.25. Assume the ψ for the new study is exchangeable with those for the previous studies.

(2)(e)(i) Re-draw your DAG, adding new nodes to represent the new $\hat{\psi}$ and new ψ



(2)(e)(ii) Correspondingly modify your JAGS model to answer the following parts. Show the modified JAGS and R code and output that you used.

```
model {
  for (j in 1:length(psi_hat)) {
    psi_hat[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

  psi_hat.tilde ~ dnorm(psi.tilde, 1/sigma.tilde^2)
  psi.tilde ~ dnorm(psi0, 1/sigmasq0)

  lead.ind = psi_hat.tilde > 2*sigma.tilde
}
```

```r
m1mod = jags.model("asgn2templatemod.bug", c(as.list(d), sigma.tilde = 0.25))
```

```
## Compiling model graph
##    Resolving undeclared variables
##    Allocating nodes
## Graph information:
##    Observed stochastic nodes: 12
##    Unobserved stochastic nodes: 16
##    Total graph size: 77
##
## Initializing model
```

(2)(e)(iii) Estimate the posterior mean and posterior standard deviation, and form a 95% central posterior predictive interval for the estimated log-odds ratio that the new study will obtain.

```r
update(m1mod, 10000)    #Burn-in

x1mod = coda.samples(m1mod, c("psi_hat.tilde", "lead.ind"), n.iter = 100000)
summary(x1mod)
```

```
##
## Iterations = 11001:111000
## Thinning interval = 1
## Number of chains = 1
## Sample size per chain = 100000
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##                Mean    SD Naive SE Time-series SE
## lead.ind      0.256 0.436  0.00138        0.00152
## psi_hat.tilde 0.214 0.482  0.00152        0.00158
##
## 2. Quantiles for each variable:
##
##                 2.5%     25%   50%   75% 97.5%
## lead.ind       0.000  0.0000 0.000 1.000   1.0
## psi_hat.tilde -0.726 -0.0885 0.206 0.508   1.2
```

posterior mean of $\tilde{\hat{\psi}} \approx$ 0.214

posterior standard deviation of $\tilde{\hat{\psi}} \approx$ 0.482

95% central posterior interval for $\tilde{\hat{\psi}} \approx$ (-0.726, 1.197)

(2)(e)(iv) Estimate the posterior predictive probability that the new estimated log-odds ratio will be at least twice its standard error, i.e., at least two standard errors (2σ) greater than zero. (This is roughly the posterior probability that the new study will find a statistically significant result, and in the positive direction.)

posterior predictive probability $Pr(\tilde{\hat{\psi}} > 2\tilde{\sigma}) \approx$ 0.256

Appendix (Data)

j	psi_hat	sigma
1	1.055	0.373
2	-0.097	0.116
3	0.626	0.229
4	0.017	0.117
5	1.068	0.471
6	-0.025	0.120
7	-0.117	0.220
8	-0.381	0.239
9	0.507	0.186
10	0.000	0.328
11	0.385	0.206
12	0.405	0.254