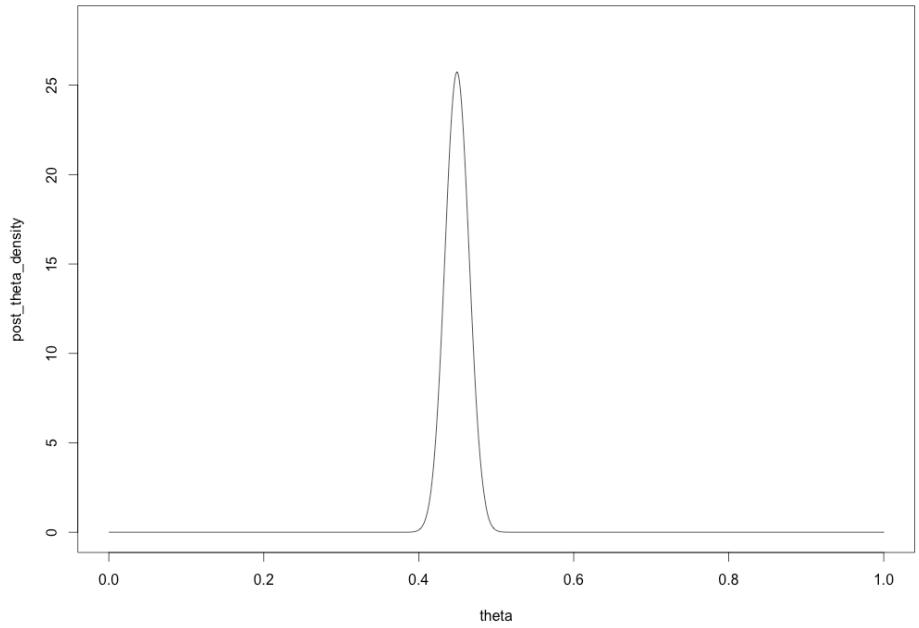
3. (a)



(b).
The posterior mean is 0.4495196 and the posterior variance is 0.0002376864

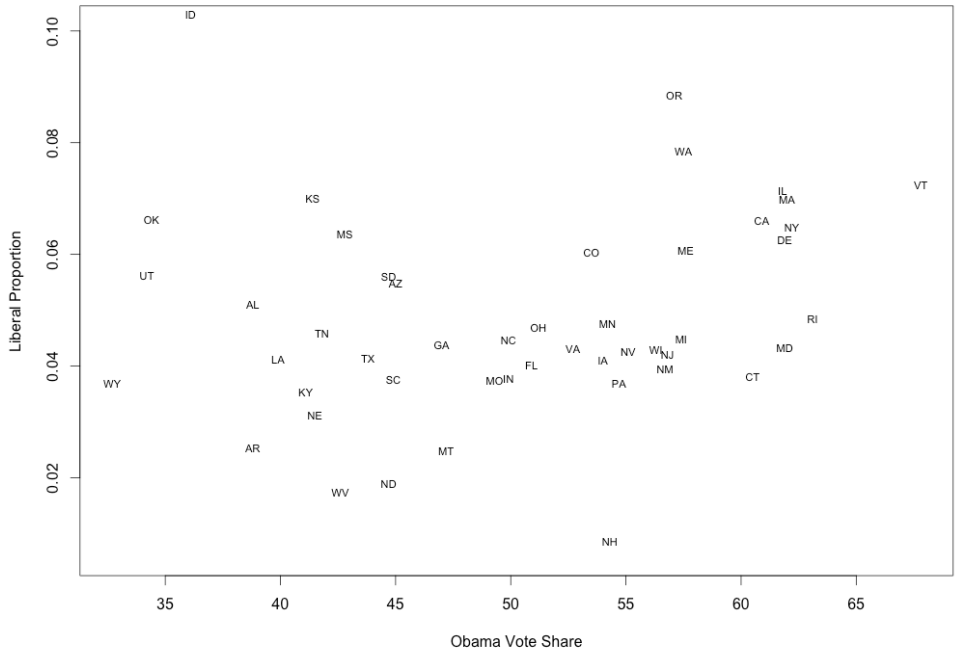
(c).
The posterior median is 0.449

(d).
The 95% central posterior interval is (0.419, 0.48).

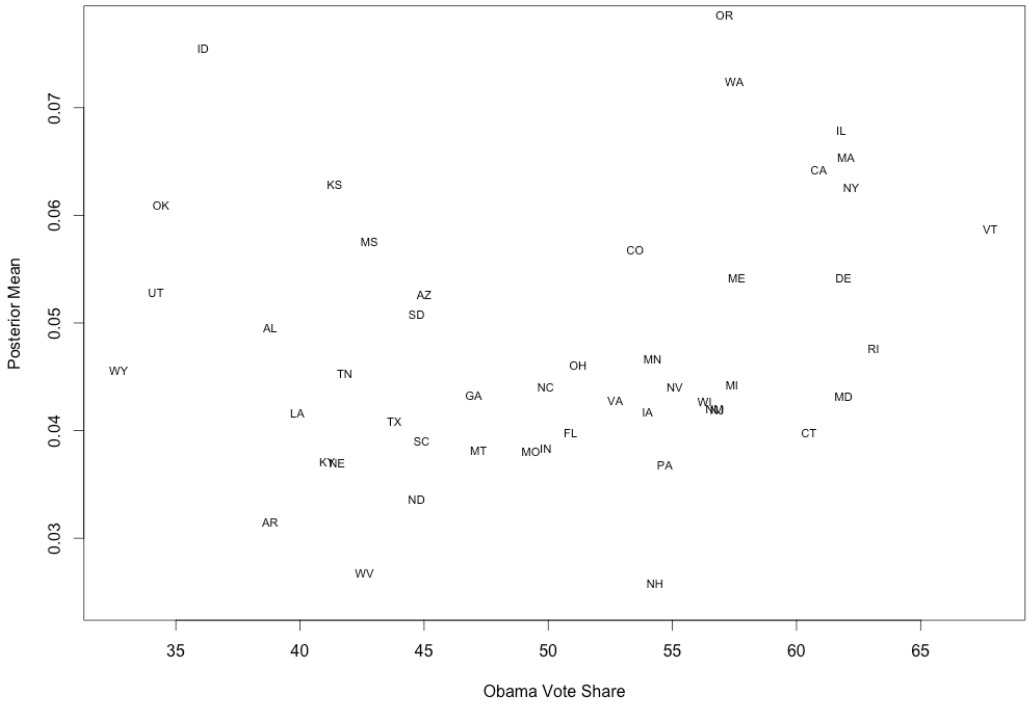
```
library(triangle)

# The value of theta
thetas <- seq(0, 1, 0.001)

# Set prior P(theta), simulate the witch's hat distribution
prior_density <- 0.5 * dunif(thetas, 0, 1) + 0.5 * dtriangle(thetas, 0.385, 0.585)
names(prior_density) <- thetas

# Plot the Prior dist of theta
plot(prior_density ~ thetas, ylim = c(0, 6), type = 'l')


# The density of y|theta
```

```
likelihood <- lapply(thetas, function(x) {dbinom(437, size = 980, prob = x)})
likelihood <- unlist(likelihood)
names(likelihood) <- thetas

# The marginal dist of y
marginal_y <- sum(prior_density * likelihood)

# Posterior dist
post_theta_prob <- prior_density * likelihood / marginal_y
post_theta_density <- post_theta_prob / (0.001 * sum(post_theta_prob))
names(post_theta_prob) <- thetas
names(post_theta_density) <- thetas

# Plot of posterior density
plot(post_theta_density ~ thetas, ylim = c(0, 1.1 * max(post_theta_density)), type = 'l')


# posterior mean and variance
mean <- sum(thetas * post_theta_prob)
variance <- sum(thetas ^ 2 * post_theta_prob) - mean ^ 2

# posterior quantile function
quantile <- function(percentage) {
  cummu_prob_1 <- 0
  cummu_prob_2 <- 0
  for (theta in thetas) {
    cummu_prob_1 <- cummu_prob_1 + post_theta_prob[as.character(theta)]
    if ((cummu_prob_1 > percentage) && (cummu_prob_2) < percentage ) {
      print(theta)
      break
    } else {
      cummu_prob_2 <- cummu_prob_2 + post_theta_prob[as.character(theta)]
    }
  }
}

# Posterior median
quantile(0.5)

# 95% central posterior interval
quantile(0.025)
quantile(0.975)
```

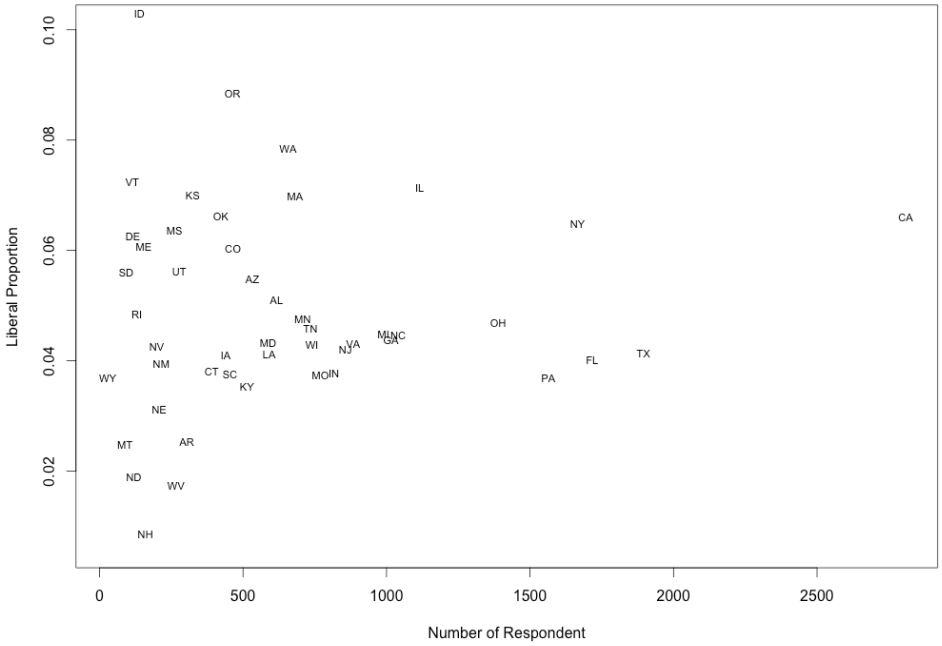
1.5(a)

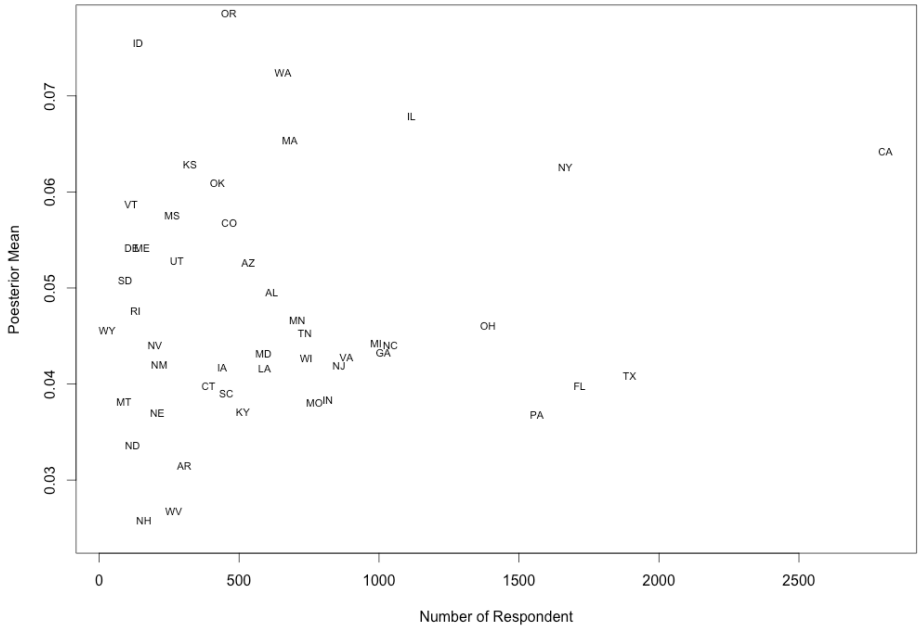

Liberal Proportion
0.02
0.04
0.06
0.08
0.10
35
40
45
50
55
60
65
Obama Vote Share
ID
OR
WA
VT
KS
IL
MA
OK
CA
NY
MS
DE
CO
ME
UT
SD
AZ
AL
RI
OH
MN
TN
GA
NC
MI
MD
VA
NV
WI
NJ
LA
TX
IA
NM
FL
CT
WY
SC
MO
IN
PA
KY
NE
AR
MT
WV
ND
NH

(b)


Posterior Mean
0.03
0.04
0.05
0.06
0.07
35
40
45
50
55
60
65
Obama Vote Share
OR
ID
WA
IL
MA
CA
KS
NY
OK
VT
MS
CO
ME
DE
UT
AZ
SD
AL
RI
MN
WY
OH
TN
NC
NV
MI
GA
VA
MD
LA
IA
TX
FL
CT
SC
MT
MO
IN
PA
NE
ND
AR
WV
NH

(c)


Liberal Proportion
0.02
0.04
0.06
0.08
0.10
0
500
1000
1500
2000
2500
Number of Respondent
ID
OR
WA
VT
KS
MA
IL
OK
NY
CA
DE
ME
MS
CO
SD
UT
AZ
AL
RI
MN
TN
OH
NV
MD
WI
VA
NJ
NC
GA
FL
TX
NM
IA
LA
CT
SC
IN
MO
PA
WY
KY
NE
MT
AR
ND
WV
NH


Poesterior Mean
0.03
0.04
0.05
0.06
0.07
0
500
1000
1500
2000
2500
Number of Respondent
OR
ID
WA
IL
MA
CA
KS
NY
OK
VT
MS
CO
ME
UT
AZ
SD
AL
RI
MN
TN
OH
WY
NV
MI
NC
GA
MD
WI
VA
NJ
NM
IA
LA
TX
CT
SC
FL
MT
IN
MO
NE
KY
PA
ND
AR
WV
NH

```
library(foreign)
library(plyr)
library(dplyr)
library(openintro)

# read data and select useful features
elect.data <- read.dta('pew_research_center_june_elect_wknd_data.dta')
Obama_votes <- read.csv('2008ElectionResult.csv')
elect.data <- elect.data[,c('ideo', 'state')]
Obama_votes <- Obama_votes[, c('state', 'vote_Obama_pct')]

# Exclude Alaska, Hawaii and Washington DC and rows with NA
elect.data <- filter(elect.data, !(state %in% c('alaska', 'washington dc', 'hawaii')))
elect.data <- elect.data[complete.cases(elect.data),]

#Transfer state name to state abbreviation
elect.data$state <- state2abbr(elect.data$state)
Obama_votes$state <- state2abbr(Obama_votes$state)

# split the election data by state, calculate the proportion of very liberal people
# and combine together

liberal_proportion <- ddply(elect.data, 'state', function(dataframe){
   liberal_people <- dataframe[dataframe$ideo == 'very liberal',]
   proportion <- nrow(liberal_people) / nrow(dataframe)
   data.frame(liberal_people = nrow(liberal_people), total_people = nrow(dataframe),
state_proportion = proportion)
})

# Combine liberal_probortion dataset and Obama_votes dataset by state abbreviation
liberal_proportion_vs_Obama_vote <- merge(liberal_proportion, Obama_votes, by = 'state')

# Plot the liberal proportion vs Obama vote
plot(liberal_proportion_vs_Obama_vote$vote_Obama_pct,
liberal_proportion_vs_Obama_vote$state_proportion,
    ylab = 'Liberal Proportion', xlab = 'Obama Vote Share', type = 'n')
text(liberal_proportion_vs_Obama_vote$vote_Obama_pct,
liberal_proportion_vs_Obama_vote$state_proportion,
    labels = liberal_proportion_vs_Obama_vote$state, cex = 0.7, pos = 3)

# Estimate the prior
prior_mean <- mean(liberal_proportion$state_proportion)
prior_var <- var(liberal_proportion$state_proportion)
```

```
alpha <- (prior_mean ^ 2) * (1 - prior_mean) / prior_var - prior_mean
beta <- alpha * (1 - prior_mean) / prior_mean

# Posterior Mean for each state
liberal_proportion_vs_Obama_vote$posterior_mean <- (alpha +
liberal_proportion$liberal_people) / (alpha + beta + liberal_proportion$total_people)

# Plot the posterior mean vs Obama vote
plot(liberal_proportion_vs_Obama_vote$vote_Obama_pct,
liberal_proportion_vs_Obama_vote$posterior_mean,
    ylab = 'Posterior Mean', xlab = 'Obama Vote Share', type = 'n')
text(liberal_proportion_vs_Obama_vote$vote_Obama_pct,
liberal_proportion_vs_Obama_vote$posterior_mean,
    labels = liberal_proportion_vs_Obama_vote$state, cex = 0.7, pos = 3)

# Plot the liberal Mean vs number of respondents
plot(liberal_proportion_vs_Obama_vote$total_people,
liberal_proportion_vs_Obama_vote$state_proportion,
    ylab = 'Liberal Proportion', xlab = 'Number of Respondent', type = 'n')
text(liberal_proportion_vs_Obama_vote$total_people,
liberal_proportion_vs_Obama_vote$state_proportion,
    labels = liberal_proportion_vs_Obama_vote$state, cex = 0.7, pos = 3)


# Plot the liberal Mean vs Number of Respondents
plot(liberal_proportion_vs_Obama_vote$total_people,
liberal_proportion_vs_Obama_vote$posterior_mean,
    ylab = 'Poesterior Mean', xlab = 'Number of Respondent', type = 'n')
text(liberal_proportion_vs_Obama_vote$total_people,
liberal_proportion_vs_Obama_vote$posterior_mean,
    labels = liberal_proportion_vs_Obama_vote$state, cex = 0.7, pos = 3)
```