Bayesian Methods - Assignment 4

Ryan Durfey

April 25, 2016

```
opar<-par(no.readonly = TRUE)
source('/Users/rdurfey/R_misc/BayesianMethods/DBDA2Eprograms/DBDA2E-utilities.R')
```

```
##
## ***********************************************************************
## Kruschke, J. K. (2015). Doing Bayesian Data Analysis, Second Edition:
## A Tutorial with R, JAGS, and Stan. Academic Press / Elsevier.
## ***********************************************************************

## Loading required package: coda

## Linked to JAGS 4.2.0

## Loaded modules: basemod,bugs

## Warning: package 'runjags' was built under R version 3.2.4
```

Part 1: Workshop

NOTE: Since there are no "skipped code" sections in the Workshop, Yuri has said that we do not need to reproduce it here (which could just be done by copy/pasting).

Part 2: Estimating the Risk of Tumor in a Group of Rats

```
Data<-c(s=14,k=4) # sample is 14 rates; 4 of them developed tumors

# use binomial model for likelihood & beta distribution for prior

# Likelihood function for calculating binomial likelihood as in the previous workshop.
likeli<-function(par,data){
  sdata<-sum(data)
  ldata<-length(data)
  return(par^sdata*(1-par)^(ldata-sdata))
}
```

Convert mean and standard deviation values into beta shape parameters

```
mu <- 0.136
sd <- 0.1034
a <- mu*(mu*(1-mu)/sd^2 - 1)
b <- (1-mu)*(mu*(1-mu)/sd^2 - 1)
c(a=a,b=b)
```

```
##        a        b
## 1.358688 8.631663
```

```
# double-check this with Kruschke's function
betaABfromMeanSD(mean=mu,sd=sd)
```

```
## $a
## [1] 1.358688
##
## $b
## [1] 8.631663
```

```
# boom.
```

Calculate parameters of posterior distribution

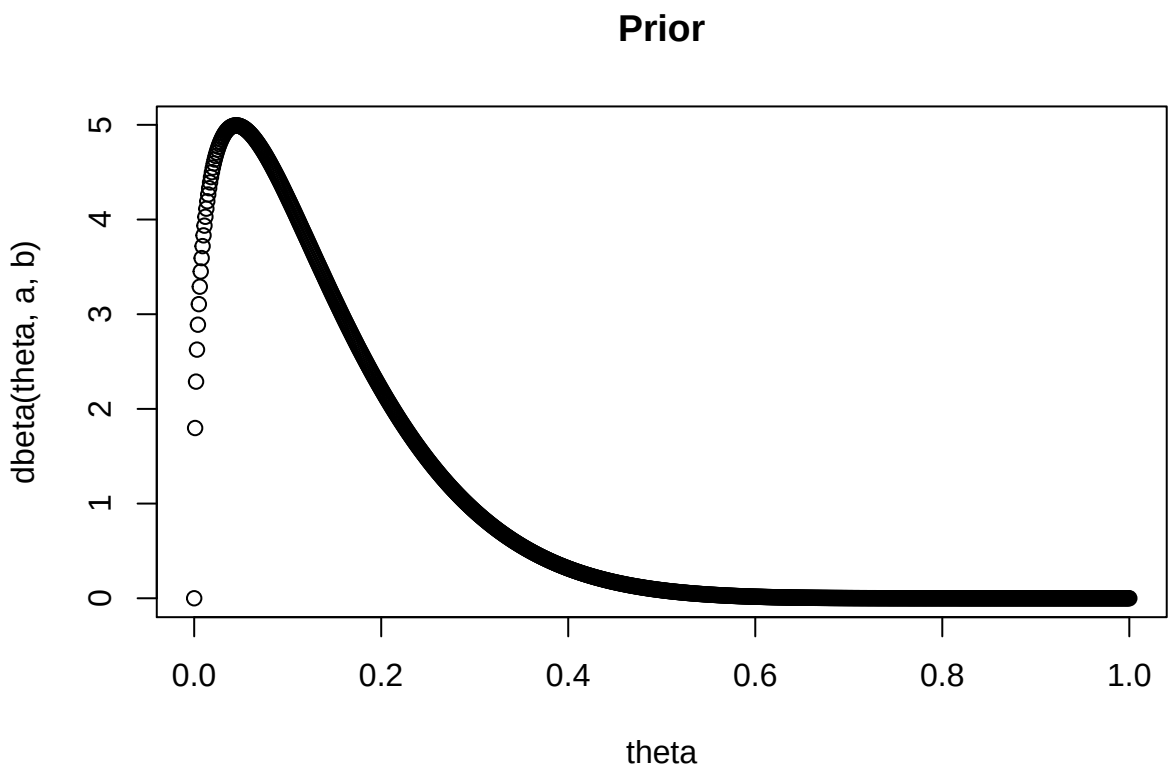
```
a.post <- a + Data[2]
b.post <- Data[1] - Data[2] + b
c(a=as.numeric(a.post),b=as.numeric(b.post))
```

```
##         a         b
##  5.358688 18.631663
```

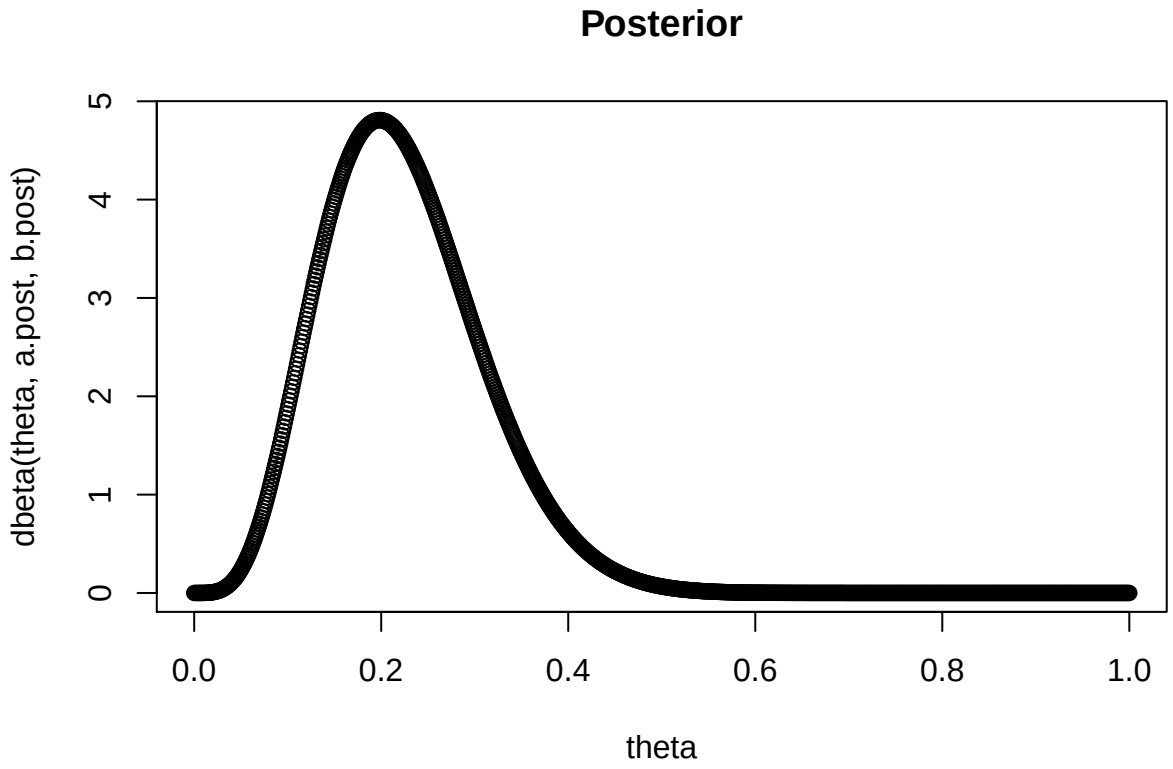
```
# here, Data[2] is z & Data[1] is N
# and the typical equations are:
## a.post = a + z
## b.post = N - z + b
# NOTE: this is for the situation where we have a Beta prior and Bernoulli likelihood
```

Compare Prior and Posterior Distributions and Interpret the Difference

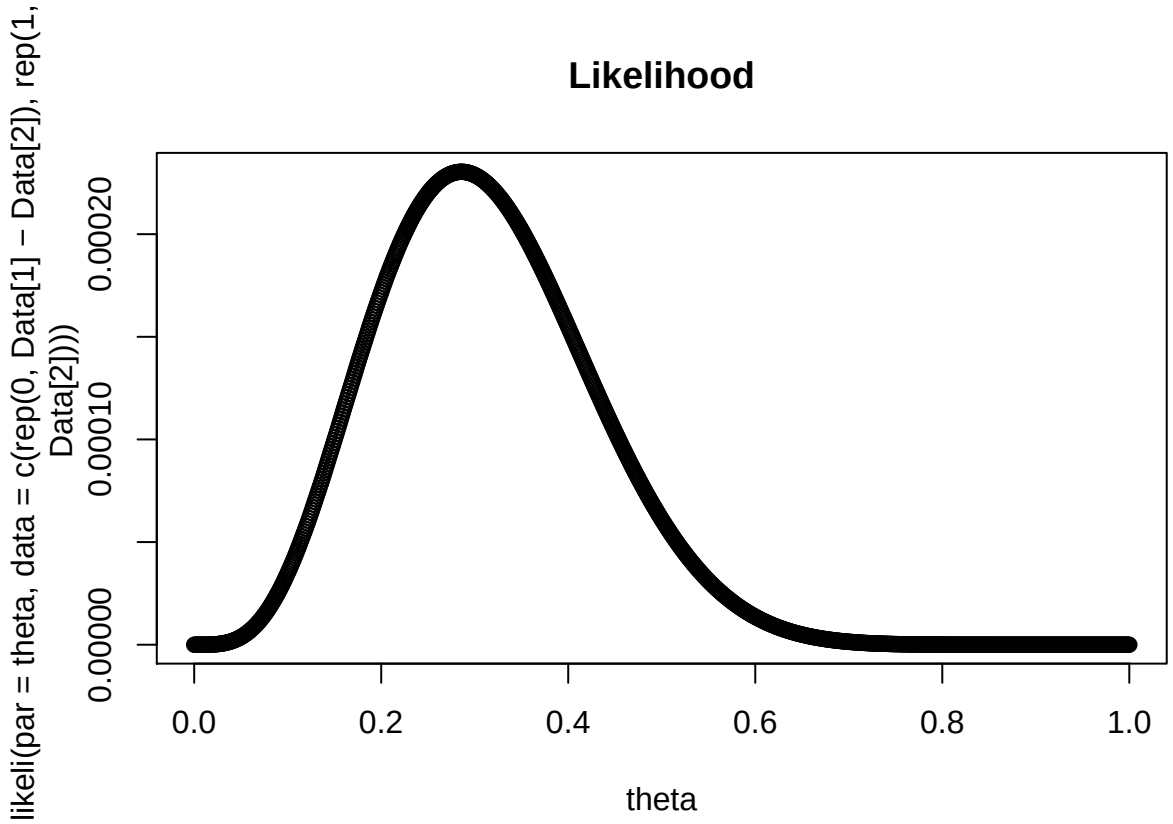
```
theta <- seq(0,1,length=1001)

# plot of prior
plot(theta,dbeta(theta,a,b),main="Prior")
```




```
# plot of posterior
plot(theta,dbeta(theta,a.post,b.post),main="Posterior")
```



```r
# plot of likelihood
plot(theta,likeli(par=theta,data=c(rep(0,Data[1]-Data[2]),rep(1,Data[2]))),main="Likelihood")
```



From the Prior and Posterior plots above, we see the Prior as heavily skewed, with the bulk of the density toward the left. The Posterior is skewed a little less, and is shifted to the right a bit. This makes intuitive sense when we also look at the Likelihood plot. This coincides with the previously observed phenomenon of the Posterior being a compromise between the Prior and Likelihood.

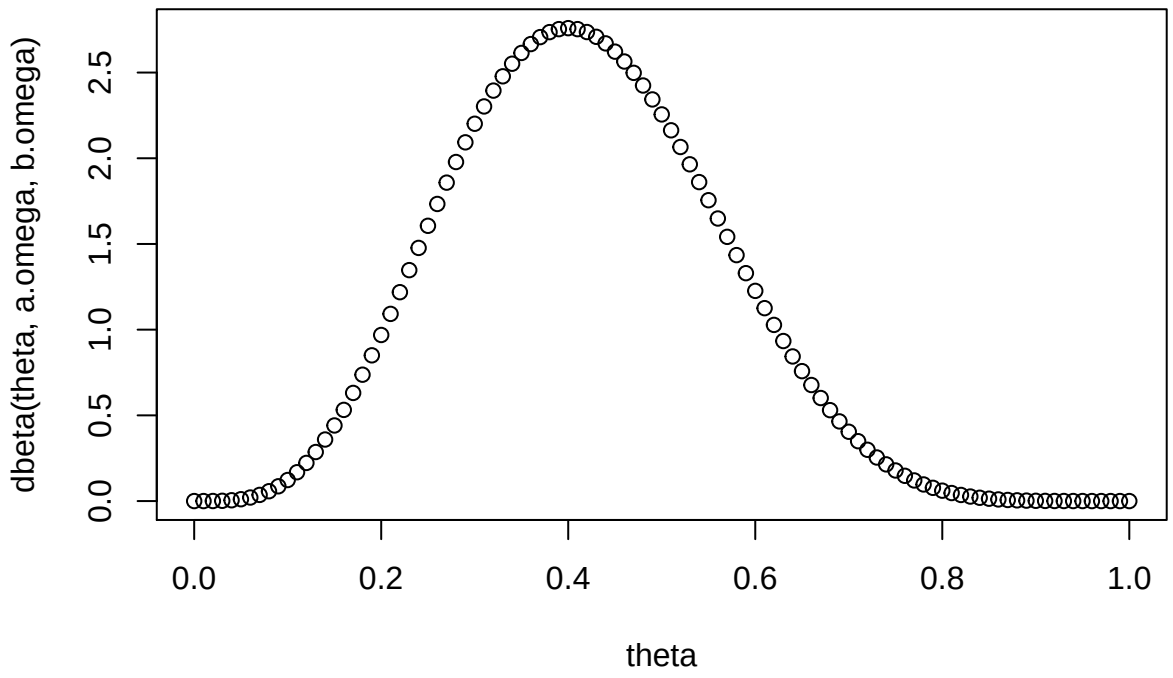
Furthermore, it's worth noting that our initial beta shape parameters for the Prior summed to approximately 10, which is the concentration. Since the data we used to update the system had 14 observations, it meant that the Likelihood had a slightly larger influence than the Prior. This is why the Posterior's mode is closer to that of the Likelihood than the Prior.

Grid Approximation to Find Posterior Distribution

```r
# Grid of Omega & Theta values
omega<-theta<-seq(0,1,length=101) # only use length=101, because 1001 takes waay too long


# Hyperprior shape parameters (belief based on historical observations)
omega.hyperprior <- 0.4 # given
kappa.hyperprior <- 12 # given
a.omega <- omega.hyperprior*(kappa.hyperprior-2)+1
b.omega <- (1-omega.hyperprior)*(kappa.hyperprior-2)+1
c(a.omega=a.omega,b.omega=b.omega)
```

```
## a.omega b.omega
##       5       7
```

```
plot(theta,dbeta(theta,a.omega,b.omega))
```

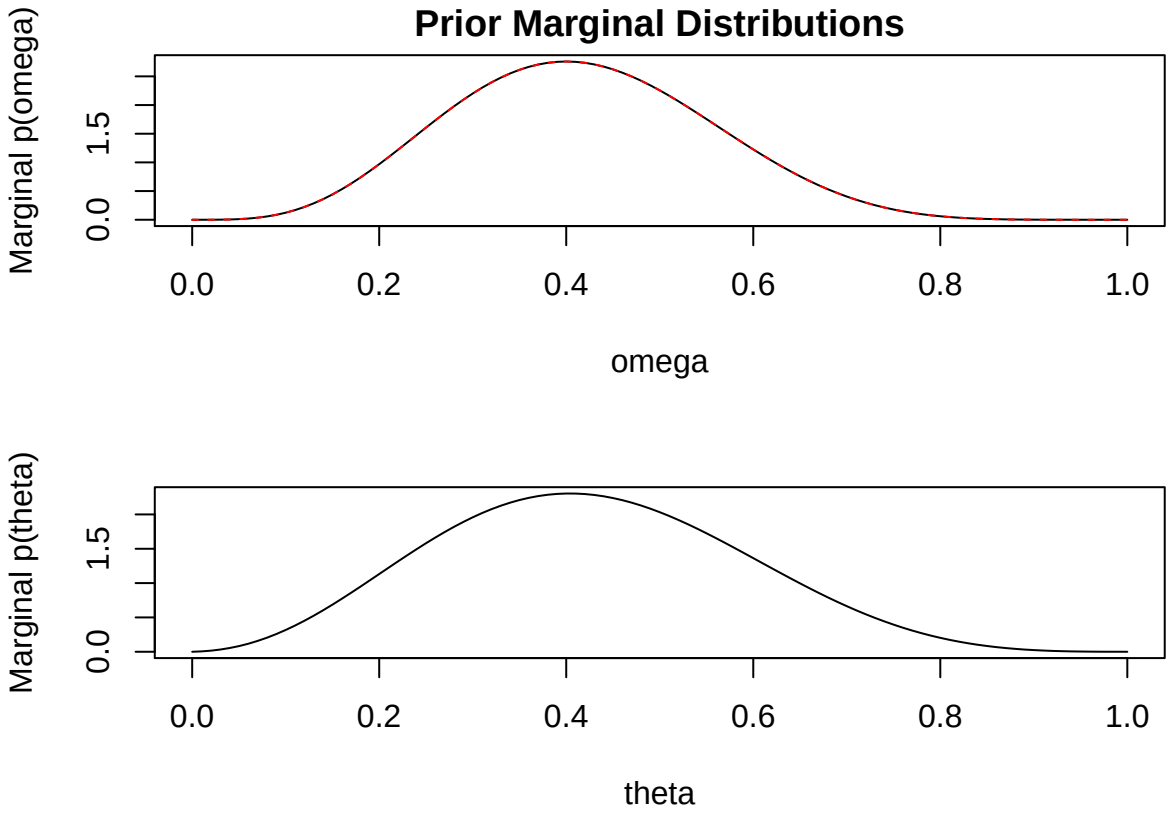



```
# confirm shape parameters with Kruschke's function
betaABfromModeKappa(mode=omega.hyperprior,kappa=kappa.hyperprior)
```

```
## $a
## [1] 5
##
## $b
## [1] 7
```

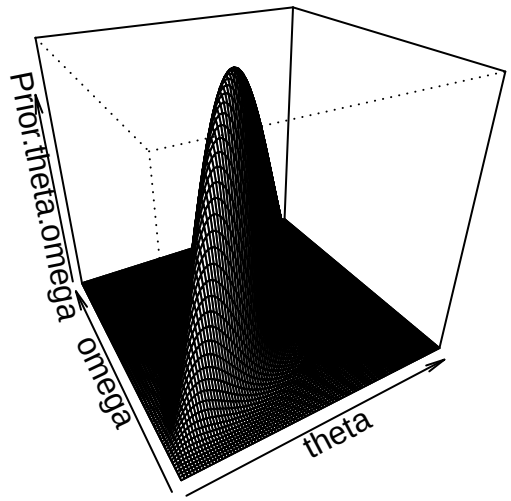
```
# boom.


# Joint Prior (omega & theta)
jointPrior<-function(theta,omega,A_omega,B_omega,K){
  res<-dbeta(omega,A_omega,B_omega)*dbeta(theta,omega*(K-2)+1,(1-omega)*(K-2)+1)
  res
}
dens <- expand.grid(omega,theta)
colnames(dens) <- c("Omega","Theta")
dens$Prior <- apply(dens,1,function(z)jointPrior(theta=z[1],omega=z[2],
                                                 A_omega=a.omega,B_omega=b.omega,K=20))
# NOTE: Kappa=20 for Prior Dist. was given in assignment

Prior.theta.omega <- matrix(dens$Prior,101,101)

Prior.omega.marginal <- apply(Prior.theta.omega,2,sum)
```

```r
Prior.omega.marginal <- Prior.omega.marginal/sum(Prior.omega.marginal)*100

Prior.theta.marginal <- apply(Prior.theta.omega,1,sum)
Prior.theta.marginal <- Prior.theta.marginal/sum(Prior.theta.marginal)*100

# plots of marginals
par(mfrow=c(2,1),mar=c(5.1,4.1,1.7,2.1))
matplot(omega,cbind(Prior.omega.marginal,dbeta(omega,a.omega,b.omega)),type='l',
        ylab='Marginal p(omega)')
title(main='Prior Marginal Distributions')
plot(theta,Prior.theta.marginal,type='l',ylab='Marginal p(theta)')
```

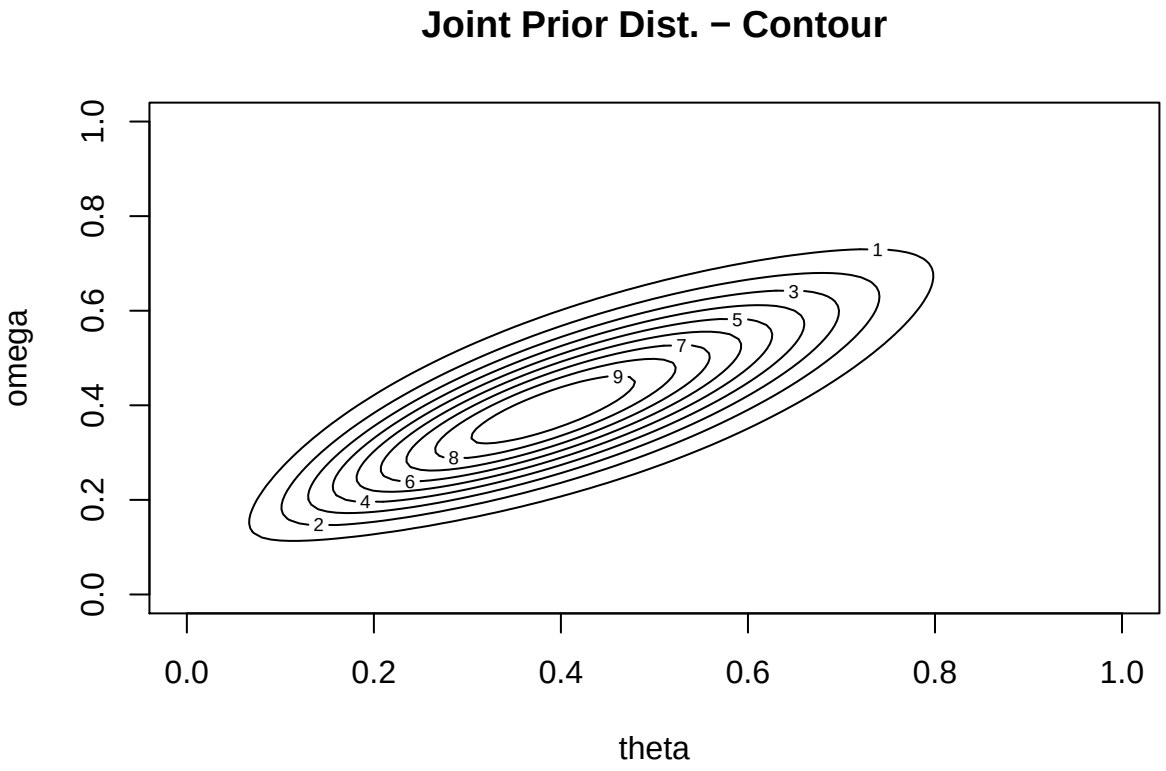



```r
par(opar)

# plots of Joint Prior
persp(theta,omega,Prior.theta.omega,d=1,theta = -30,phi=30,
      main='Joint Prior Distribution')
```

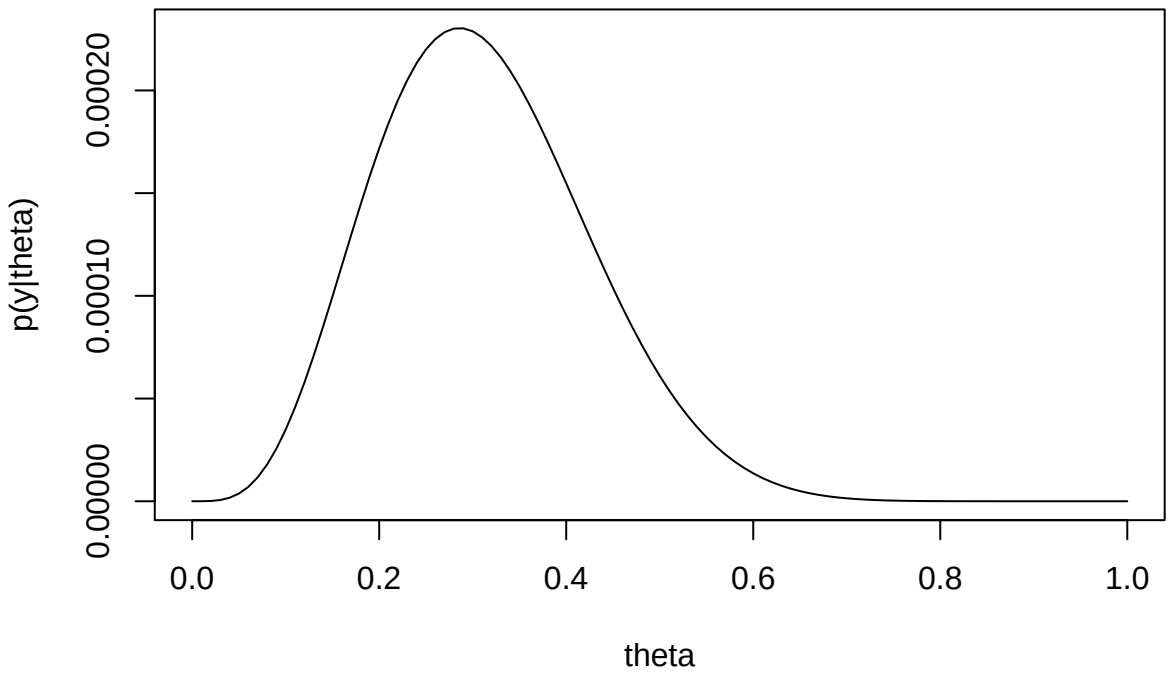
Joint Prior Distribution




```
contour(x=omega,y=theta,z=Prior.theta.omega,xlab='theta',ylab='omega',
        main="Joint Prior Dist. - Contour")
```




```
# Likelihood
like <- likeli(par=theta,data=c(rep(0,Data[1]-Data[2]),rep(1,Data[2])))
plot(theta,like,type='l',ylab='p(y|theta)',main="Likelihood")
```

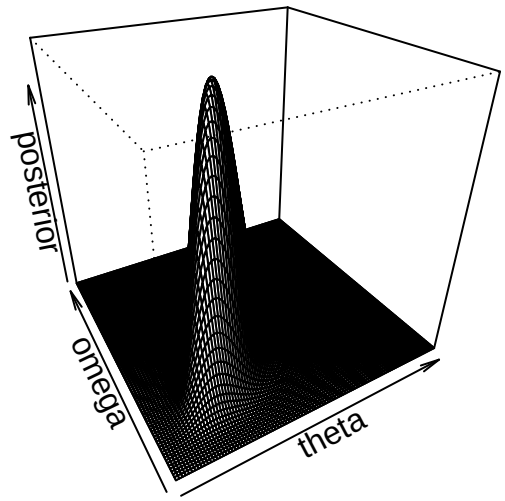



```
# Posterior
posterior <- apply(Prior.theta.omega,2,function(z)z*like)
posterior <- posterior/sum(posterior)
posterior[1:5,1:5]
```

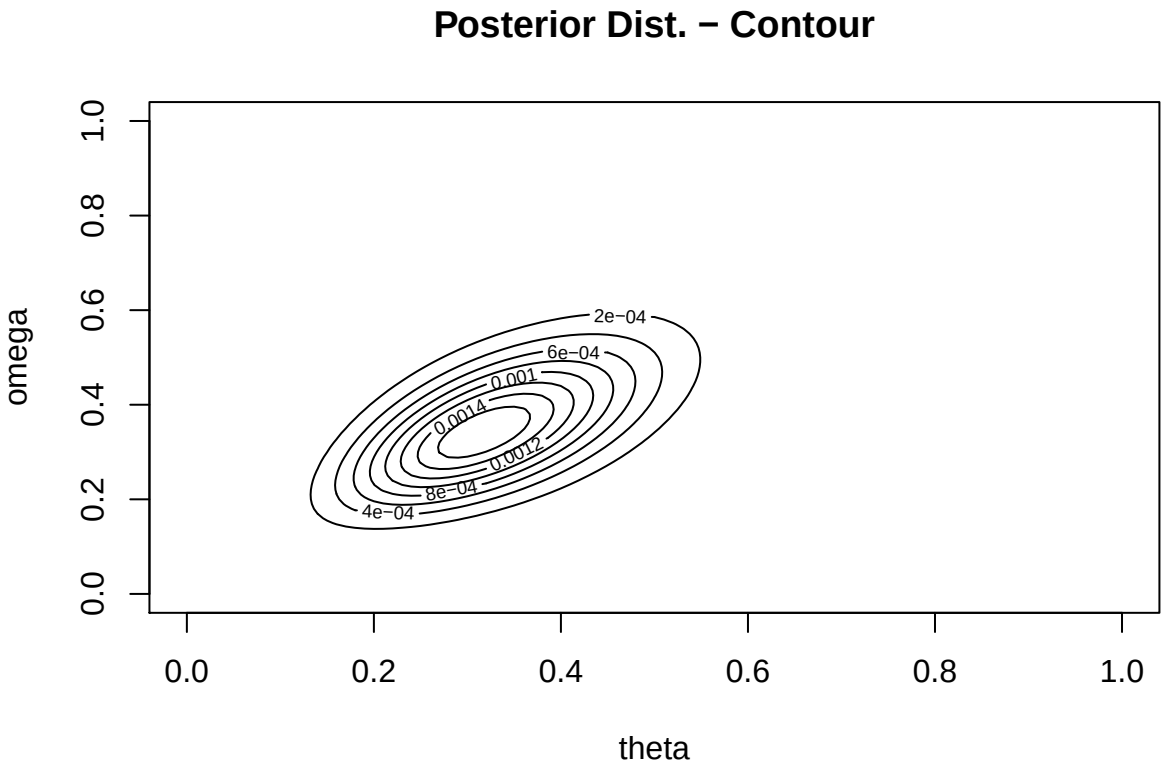
```
##      [,1]         [,2]         [,3]         [,4]         [,5]
## [1,]    0 0.000000e+00 0.000000e+00 0.000000e+00 0.000000e+00
## [2,]    0 2.161471e-12 2.490093e-11 8.745727e-11 1.857609e-10
## [3,]    0 2.953882e-11 3.862231e-10 1.539564e-09 3.711378e-09
## [4,]    0 1.209302e-10 1.704033e-09 7.320416e-09 1.901829e-08
## [5,]    0 3.016997e-10 4.485565e-09 2.033171e-08 5.573252e-08
```

```
# plots of Posterior
persp(theta,omega,posterior,d=1,theta = -30,phi=30,
      main='Posterior Distribution')
```

Posterior Distribution

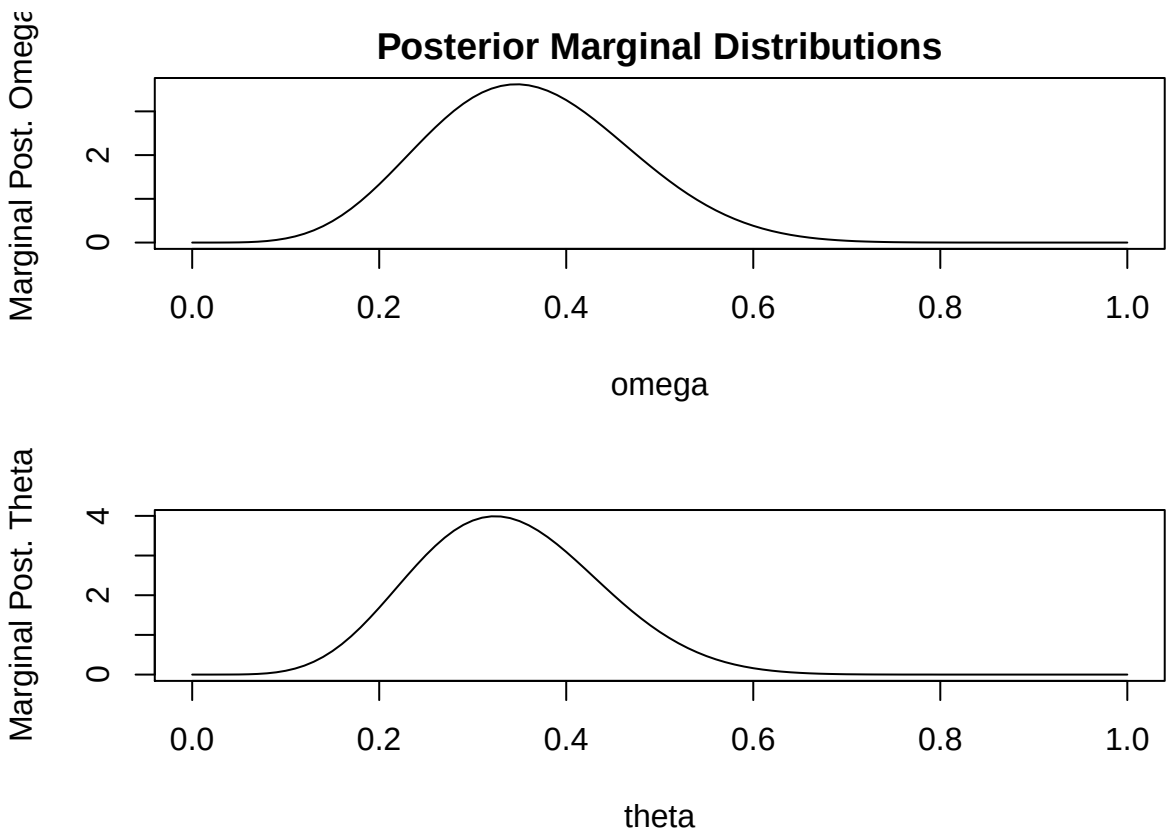



```
contour(x=omega,y=theta,z=posterior,xlab='theta',ylab='omega',
        main="Posterior Dist. - Contour")
```




```
# Posterior Marginal Dists
Posterior.omega.marginal<-apply(posterior,2,sum)
Posterior.omega.marginal<-Posterior.omega.marginal/sum(Posterior.omega.marginal)*100

Posterior.theta.marginal<-apply(posterior,1,sum)
Posterior.theta.marginal<-Posterior.theta.marginal/sum(Posterior.theta.marginal)*100
```

```r
# plots of posterior marginals
par(mfrow=c(2,1),mar=c(5.1,4.1,1.7,2.1))
plot(omega,Posterior.omega.marginal,type="l",ylab='Marginal Post. Omega')
title(main='Posterior Marginal Distributions')
plot(theta,Posterior.theta.marginal,type="l",ylab='Marginal Post. Theta')
```




```r
par(opar)
```