

# Xinshi-Ma-HW2

Jan. 10, 2022

## LAB

Observed data:

```
# set a seed
set.seed(123)
# create the observed data
obs.data <- rbinom(n = 100, size = 1, prob = 0.01)
# inspect the observed data
# head(obs.data)
# tail(obs.data)
# length(obs.data)
```

## Task 3

Write a function that takes as its inputs that data you simulated (or any data of the same type) and a sequence of  $\theta$  values of length 1000 and produces Likelihood values based on the Binomial Likelihood. Plot your sequence and its corresponding Likelihood function.

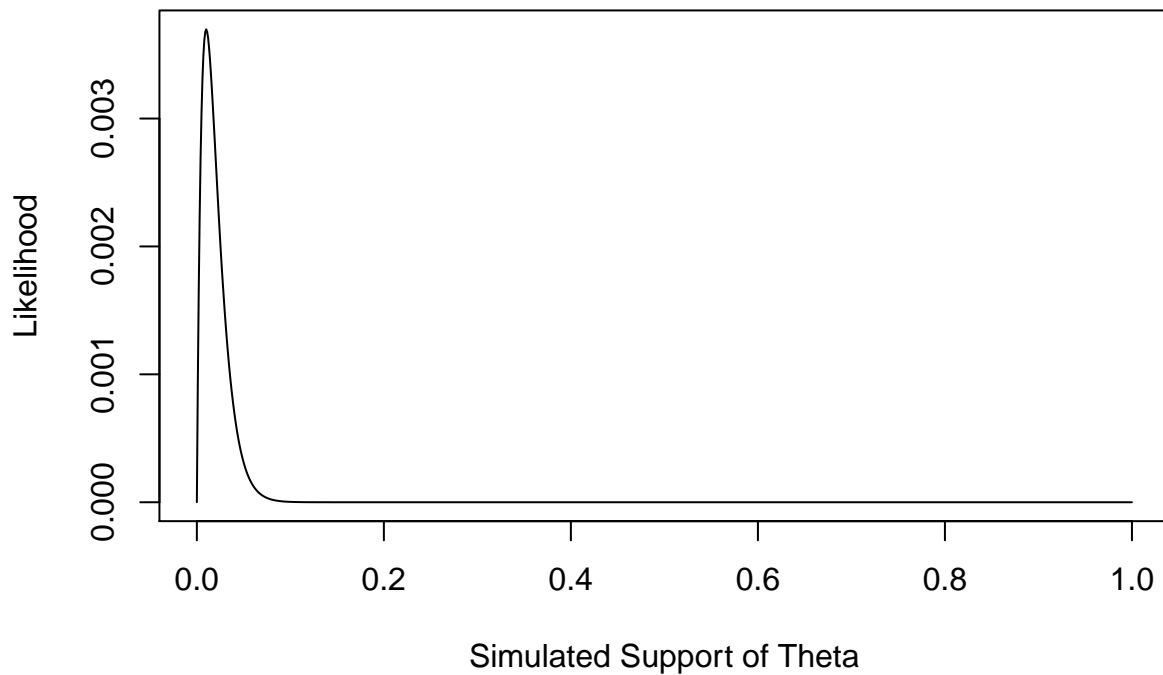
The likelihood function is given below. Since this is a probability and is only valid over the interval from  $[0, 1]$  we generate a sequence over that interval of length 1000.

You have a rough sketch of what you should do for this part of the assignment. Try this out in lab on your own.

```
### Bernoulli LH Function ###
# Input: obs.data, theta
# Output: bernoulli likelihood
LH.func = function(obs.data, theta){
  x = sum(obs.data)
  N = length(obs.data)
  LH = (theta^x)*(1-theta)^(N-x)
  return(LH)
}

### Plot LH for a grid of theta values ###
# Create the grid #
theta.sim = seq(0,1,length.out=1000)
# Store the LH values
LH.sim = LH.func(obs.data, theta.sim)
# Create the Plot
plot(x=theta.sim, y=LH.sim, type='l', main='Likelihood Profile',
     xlab="Simulated Support of Theta", ylab='Likelihood')
```

## Likelihood Profile



### Task 4 (To be completed for homework)

Write a function that takes as its inputs prior parameters  $a$  and  $b$  for the Beta-Bernoulli model and the observed data, and produces the posterior parameters you need for the model. **Generate and print** the posterior parameters for a non-informative prior i.e.  $(a,b) = (1,1)$  and for an informative case  $(a,b) = (3,1)$ .

```
#Function for posterior parameters
PostPram.func = function(a,b,obs.data){
  x = sum(obs.data)
  N = length(obs.data)
  Post.a = x + a
  Post.b = N - x + b
  ret = list("a" = Post.a, "b" = Post.b)
  return(ret)
}

##Generate and Print Posterior Parameters
#Case 1: non-informative prior
Pram.NonInf = PostPram.func(1,1,obs.data)
Pram.NonInf

## $a
## [1] 2
##
## $b
```

```
## [1] 100
```

Posterior Parameters for non-informative prior:  $a = 2$ ,  $b = 100$

```
#Case 2: informative prior
Pram.Info = PostPram.func(3,1,obs.data)
Pram.Info
```

```
## $a
## [1] 4
##
## $b
## [1] 100
```

Posterior Parameters for informative prior:  $a = 4$ ,  $b = 100$

## Task 5 (To be completed for homework)

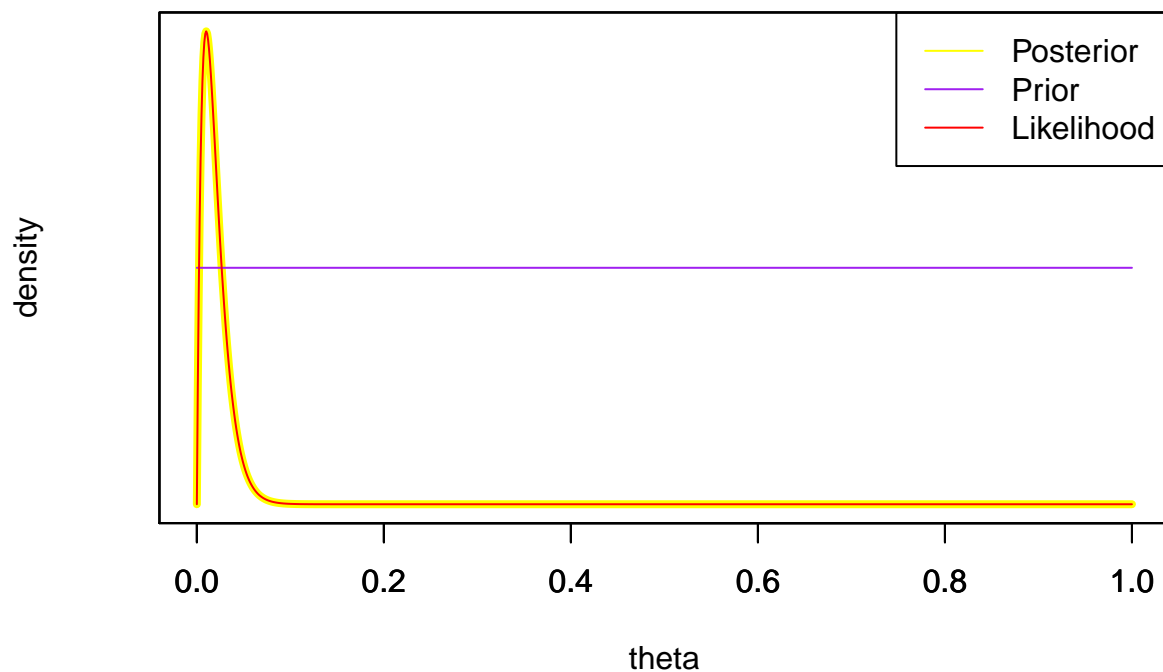
Create two plots, one for the informative and one for the non-informative case to show the posterior distribution and superimpose the prior distributions on each along with the likelihood. What do you see? Remember to turn the y-axis ticks off since superimposing may make the scale non-sense.

```
#create support
p = seq(0,1, length=1000)

##plot Non-Informative Prior and Posterior along with likelihood
#plot posterior
plot(p, dbeta(p, Pram.NonInf$a,Pram.NonInf$b),lwd=4, ylab='density',
      type = 'l', col='yellow', main = "Non-Informative Prior", xlab = "theta", yaxt='n')
par(new=TRUE)
#plot prior
plot(p, dbeta(p, 1, 1), col='purple', type = 'l',ylab = "",yaxt='n',xlab="")
par(new=TRUE)
#plot likelihood
plot(p, LH.sim, col='red', type = 'l', ylab = "",yaxt='n',xlab="")

#add legend
legend("topright", c( "Posterior","Prior",'Likelihood'),
      lty=c(1,1,1),col=c("yellow","purple","red"))
```

## Non-Informative Prior

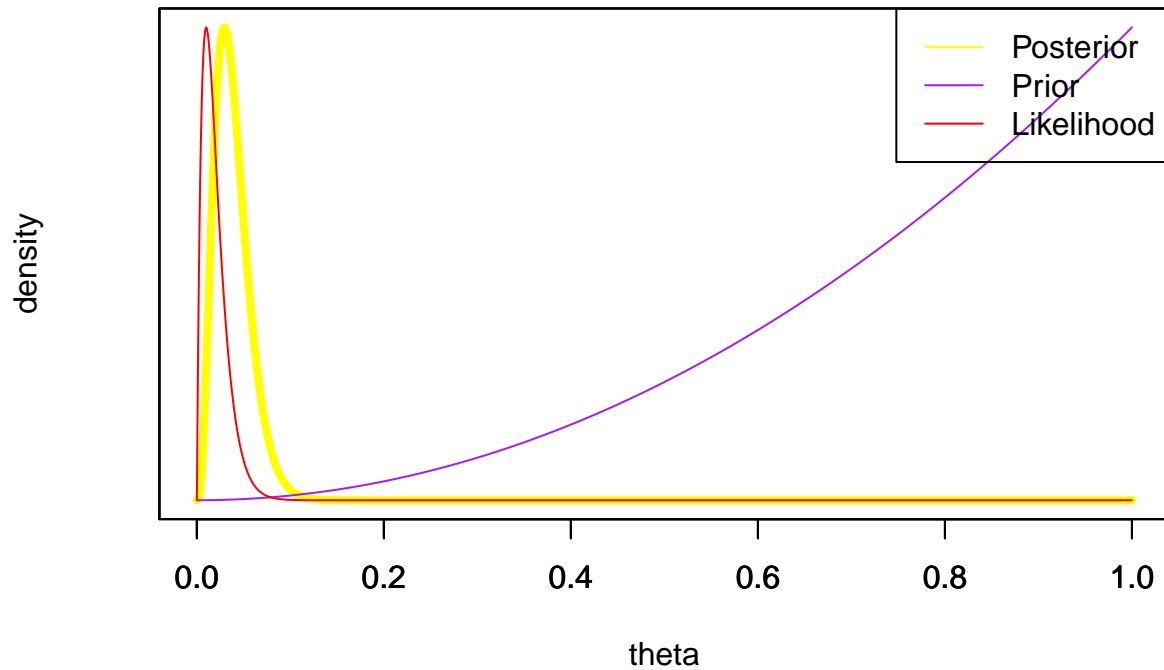


```
#create support
p = seq(0,1, length=1000)

#plot Informative Prior and Posterior along with likelihood
plot(p, dbeta(p, Pram.Info$a,Pram.Info$b),lwd=4, ylab='density',
      type='l', col='yellow', main = "Informative Prior", xlab = "theta", yaxt='n')
par(new=TRUE)
#plot prior
plot(p, dbeta(p, 3, 1), col='purple', type = 'l',ylab = "",yaxt='n',xlab="")
par(new=TRUE)
#plot likelihood
plot(p, LH.sim, col='red', type = 'l', ylab = "",yaxt='n',xlab="")

#add legend
legend("topright", c( "Posterior","Prior",'Likelihood'),lty=c(1,1,1),
      col=c("yellow","purple","red"))
```

## Informative Prior



Observation: The non-informative prior produced a posterior that is the same as the likelihood while the skewed informative prior created a posterior that is different from the likelihood, in the same direction as the prior's skewness. This is due to the fact that with an uninformative prior, our best estimation is still the likelihood function (the posterior).

2

A, B

2.

$$(a) \quad p(\theta | x_{1:n}) \propto p(x_{1:n} | \theta) \cdot p(\theta)$$

$$p(x_{1:n} | \theta) = \prod_{i=1}^n \theta \exp(-\theta x_i) \\ = \theta^n e^{-\theta \sum_{i=1}^n x_i}$$

$$p(x_{1:n} | \theta) \cdot p(\theta) = \theta^n e^{-\theta \sum_{i=1}^n x_i} \cdot \frac{b^a}{\Gamma(a)} \theta^{a-1} e^{-b\theta} \\ \propto \theta^{n+a-1} e^{-\theta(b + \sum x_i)}$$

$$\propto \text{Gamma}(n+a, b + \sum x_i)$$

(b) Since the p.d.f. integrates to 1 given it is a Gamma distribution, meaning it is a proper p.d.f.

B continued:

Condition: (Given that both parameters of the Gamma distribution are greater than 0), the posterior distribution integrates to 1.

C

```
#define data and support
obs.data = c(20.9, 69.7, 3.6, 21.8, 21.4, 0.4, 6.7, 10.0)
theta.sim = seq(0,1, length=1000)
```

```

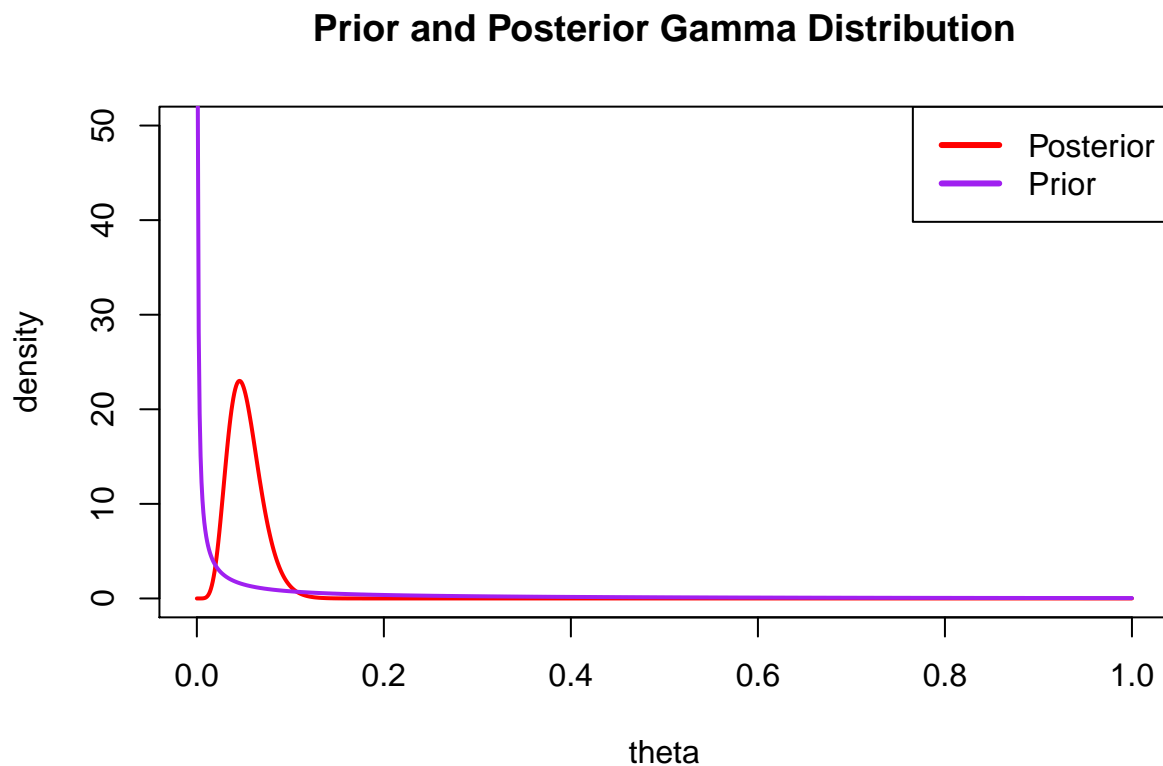
#plotting function
plot.func = function(a,b,theta,obs.data){
  x = sum(obs.data) #get x (sum of observed)
  n = length(obs.data) # get n (number of observed)
  prior = dgamma(theta,a,b)
  posterior = dgamma(theta,n+a,b+x)

#plot posterior
plot(theta,posterior,lwd=2, ylab='density', type='l', col='red',
      main = "Prior and Posterior Gamma Distribution", xlab = "theta", ylim = c(0,50))
#plot prior
lines(theta, prior,lwd=2, col='purple', type='l', ylab="", xlab="")

legend("topright", c("Posterior","Prior"), lty=c(1,1), lwd=c(3,3),
      col = c("red", "purple"))
}

plot.func(0.1, 1.0, theta.sim, obs.data)

```



## D

An appropriate way to apply the exponential distribution is to model occurrences of hurricanes. If we can roughly approximate the occurrences of hurricanes as a process in which events occur continuously and independently at a constant average rate, then we can use the exponential model to predict time between hurricanes.

Exponential distributions should not be used to predict the proportion of votes a certain candidate would

receive/probability of the candidate winning or the number of heads that come up after flipping a coin certain number of times as they are binary outcomes that are more appropriately modeled with Binomial/Bernoulli distributions.

### 3

#### A

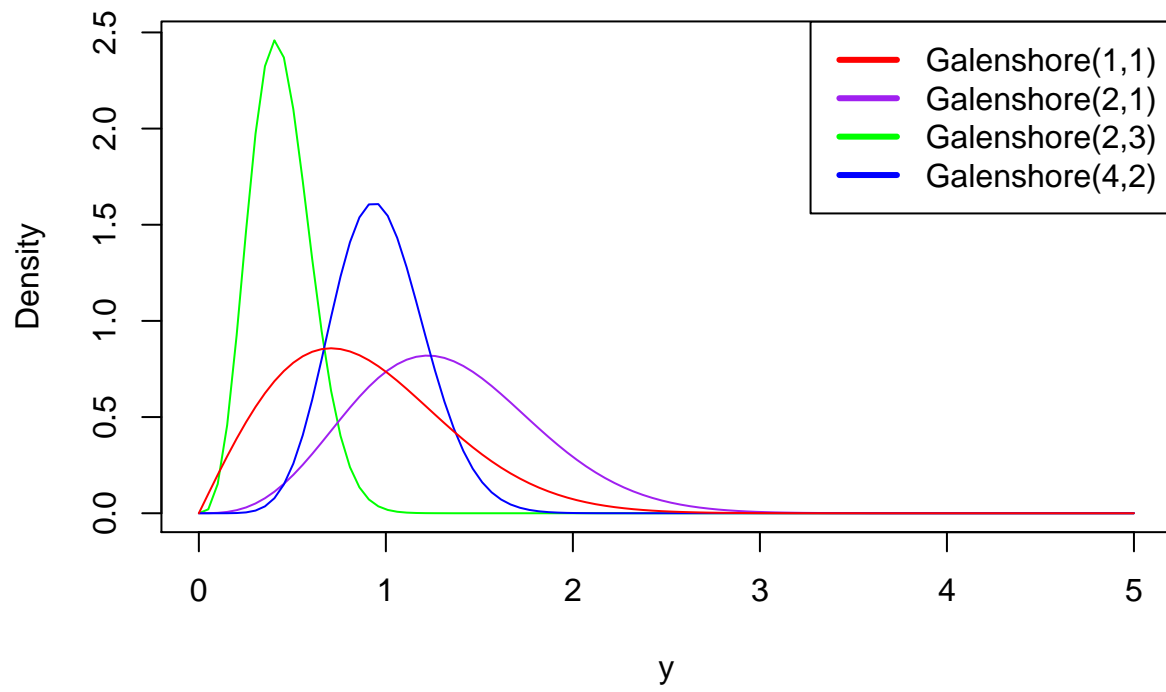
A class of conjugate prior densities for  $\theta$  also belong to the Galenshore family. Given that prior parameters are (c,d): (1,1), (2,1), (2,3), (4,2)

```
galenshore.func <- function(a, theta, y){
  p <- 2/gamma(a)*theta^(2*a)*y^(2*a-1)*exp(-theta^2*y^2)
  return(p)
}
y <- seq(0, 5, length.out = 100)
first <- galenshore.func(1,1, y)
second <- galenshore.func(2,1,y)
third <- galenshore.func(2,3,y) # type = "galenshore(2,3)"
fourth <- galenshore.func(4,2,y) #type = "galenshore(4,2)"

plot(y, third, type = "l", xlab = "y", ylab = "Density",
      main = "Galenshore Priors",col = "green")
lines(y, second, col = "purple")
lines(y, fourth, col = "blue")
lines(y, first, col = "red")
legend("topright", c( "Galenshore(1,1)","Galenshore(2,1)","Galenshore(2,3)","Galenshore(4,2)"),
      lty=c(1,1), lwd=c(3,3),
      col = c("red", "purple","green","blue"))
```



## Galenshore Priors



B

3. (b)

$$p(\theta | y_{1:n}) \propto p(y_{1:n} | \theta) p(\theta)$$

Likelihood:

$$p(y_{1:n} | \theta) = \prod_{i=1}^n \frac{2}{\Gamma(n)} \theta^{2n} y_i^{2n-1} e^{-\theta^2 y_i^2} \\ \propto \theta^{2nn} e^{-\theta^2 \sum y_i^2}$$

Prior distributed as

$$\sim \text{Gamma}(c, d) : p(\theta | c, d) = \frac{2}{\Gamma(c)} \theta^{2c-1} e^{-d\theta^2} \\ \propto \theta^{2c-1} e^{-d\theta^2}$$

$$p(y_{1:n} | \theta) p(\theta) \propto \theta^{2c-1} e^{-d\theta^2} \theta^{2nn} e^{-\theta^2 \sum y_i^2}$$

$$= \theta^{2(\text{nat } c)-1} e^{-\theta^2 (d^2 + \sum y_i^2)}$$

$$\propto \text{Gamma}(\text{nat } c, \sqrt{d^2 + \sum y_i^2})$$

C,D

(c)

$$\frac{p(\theta_a | y_{1:n})}{p(\theta_b | y_{1:n})} = \frac{\frac{2}{\Gamma(natc)} \theta_a^{2(natc)-1} e^{-(d^2 + \sum y_i^2) \theta_a^2} (\sqrt{d^2 + \sum y_i^2})^{2(natc)}}{\frac{2}{\Gamma(natc)} \theta_b^{2(natc)-1} e^{-(d^2 + \sum y_i^2) \theta_b^2} (\sqrt{d^2 + \sum y_i^2})^{2(natc)}}$$

$$= \left( \frac{\theta_a}{\theta_b} \right)^{2(natc)-1} e^{-(\theta_a^2 - \theta_b^2)(d^2 + \sum y_i^2)}$$

$$= \left( \frac{\theta_a}{\theta_b} \right)^{2(natc)-1} e^{(\theta_b^2 - \theta_a^2)(d^2 + \sum y_i^2)}$$

$\sum_{i=1}^n y_i^2$  is a sufficient statistic

(d)  $E[\theta | y_{1:n}]$

$$E[Y] = \frac{\Gamma(a + \frac{1}{2})}{\theta \Gamma(a)} \quad \text{for } p(y|\theta) \sim \text{Galensterc}(a, \theta)$$

$$E(\theta | y_{1:n}) = \frac{\Gamma(natc + \frac{1}{2})}{\sqrt{\sum y_i^2 + d^2} \Gamma(natc)}$$

E

$$(e) \quad p(y_{n+1} | y_{1:n})$$

$$= \int p(y_{n+1}, \theta | y_{1:n}) d\theta$$

$$= \int p(y_{n+1} | \theta) p(\theta | y_{1:n}) d\theta$$

$$= \int \frac{2}{\Gamma(\alpha)} \theta^{2\alpha} y_{n+1}^{2\alpha-1} e^{-\theta^2 y_{n+1}^2} \cdot \text{Galenshore}(\theta | \text{nat } c, \sqrt{d^2 + \sum y_i^2}) d\theta$$

$$= \int \frac{2}{\Gamma(\alpha)} \theta^{2\alpha} y_{n+1}^{2\alpha-1} e^{-\theta^2 y_{n+1}^2} \frac{2}{\Gamma(\text{nat } c)} \theta^{2(\text{nat } c)-1} e^{-(d^2 + \sum y_i^2)} \theta^2 \left( \sqrt{d^2 + \sum y_i^2} \right)^{2(\text{nat } c)} d\theta$$

$$= \frac{2}{\Gamma(\alpha)} \cdot \frac{2}{\Gamma(\text{nat } c)} y_{n+1}^{2\alpha-1} (d^2 + \sum y_i^2)^{(\text{nat } c)} \int \theta^{2\alpha} e^{-\theta^2 y_{n+1}^2} \theta^{2(\text{nat } c)-1} e^{-(d^2 + \sum y_i^2)} \theta^2 d\theta$$

$$= \frac{2}{\Gamma(\alpha)} \cdot \frac{2}{\Gamma(\text{nat } c)} y_{n+1}^{2\alpha-1} (d^2 + \sum y_i^2)^{(\text{nat } c)} \int \underbrace{\theta^{2(\text{nat } c + \alpha)} e^{-(y_{n+1}^2 + d^2 + \sum y_i^2)} \theta^2}_{\substack{\text{Galenshore without normalizing} \\ \text{constant}}} d\theta$$

$$= \frac{4}{\Gamma(\alpha) \Gamma(\text{nat } c)} \cdot \frac{\Gamma(\alpha + \text{nat } c)}{2} \cdot y_{n+1}^{2\alpha-1} (d^2 + \sum y_i^2)^{(\text{nat } c)} \cdot (y_{n+1}^2 + d^2 + \sum y_i^2)^{-(\text{nat } c + \alpha)}$$

$$\cdot \int \text{Galenshore}(\text{nat } c + \alpha, y_{n+1}^2 + d^2 + \sum y_i^2) d\theta$$

$$= \frac{2 \Gamma(\alpha + \text{nat } c)}{\Gamma(\alpha) \Gamma(\text{nat } c)} \cdot \frac{y_{n+1}^{2\alpha-1} (d^2 + \sum y_i^2)^{(\text{nat } c)}}{(y_{n+1}^2 + d^2 + \sum y_i^2)^{(\text{nat } c + \alpha)}}$$