

Homework 02

Christopher Cameron

Due 9/4/2021

1. Please refer to lab 2 and complete tasks 3—5.

```
# 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)

## [1] 0 0 0 0 0 0

tail(obs.data)

## [1] 0 0 0 0 0 0

length(obs.data)

## [1] 100
```

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 θ 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

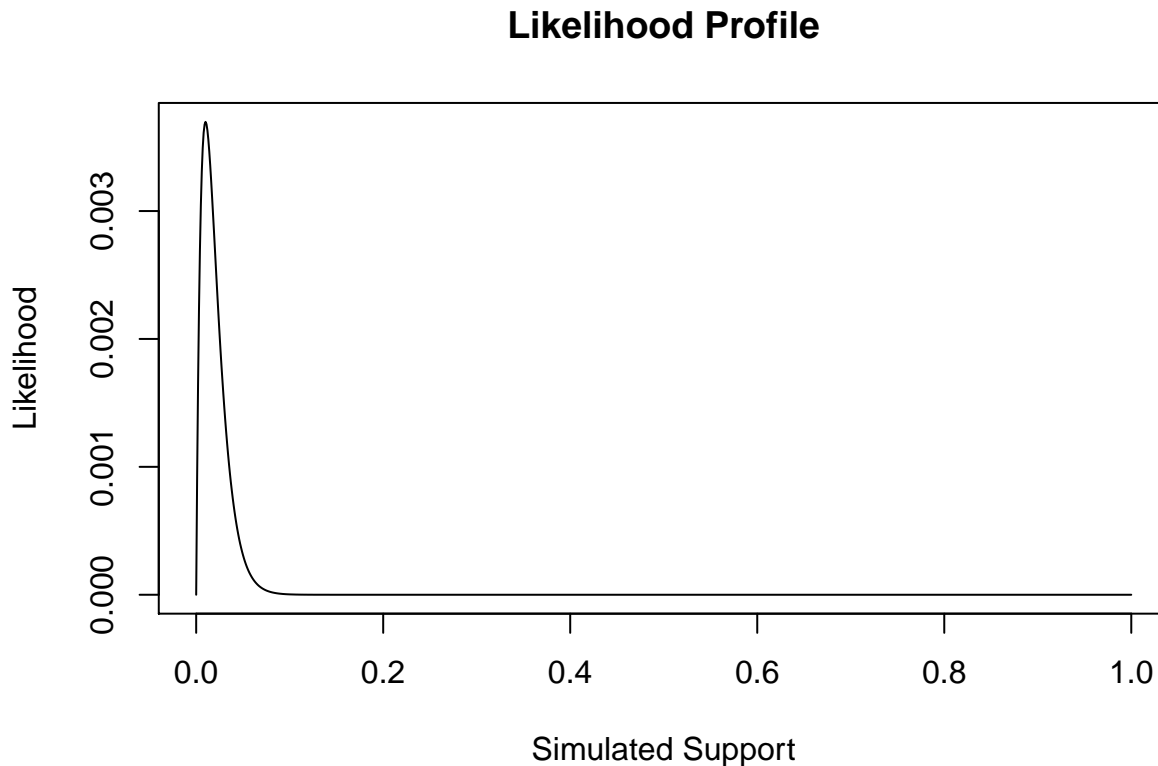
bernoulliLH <- function(obs.data, theta){
  N <- length(obs.data)
  x <- sum (obs.data)
  LH <- (theta^x)*((1-theta)^(N-x))
  return(LH)
}

### Plot LH for a grid of theta values ###
# Create the grid #
theta.sim <- seq(from = 0, to = 1, length.out = 1000)
```

```

# Store the LH values
sim.LH <- bernoulliLH(obs.data, theta = theta.sim)
# Create the Plot
plot(theta.sim, sim.LH, type = "l", main = "Likelihood Profile",
      xlab = "Simulated Support", ylab = "Likelihood")

```



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 to return parameters for beta-bernoulli posterior parameters

```

```

betberPos <- function(a, b, obs.data){
  n <- length(obs.data)
  x <- sum (obs.data)

  prior <- c((x+a), (n-x+b))
  return( prior)
}

```

```

#Print results for each case
print("Non-Informative:")

```

```

## [1] "Non-Informative:"
print(betberPos(1,1,obs.data))

```

```

## [1] 2 100

```

```
print("Informative:")

## [1] "Informative:"
print(betberPos(3,1,obs.data))

## [1] 4 100
```

Task 5 (To be completed for homework)

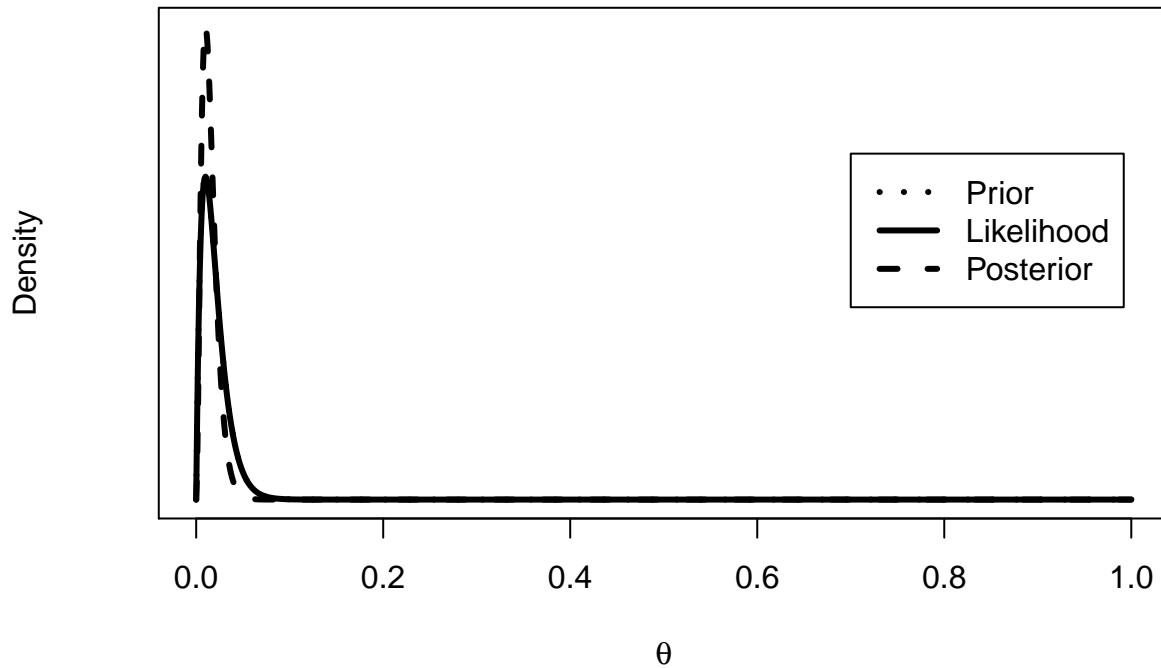
Create two plots, one for the informative and one for the non-informative case to show the posterior distribution and superimpose the prior distributions on each along with the likelihood. What do you see? Remember to turn the y-axis ticks off since superimposing may make the scale non-sense.

```
#Function to create plot of beta distribution
#p1 p2 - parameters
distPlot <- function(p1,p2){
  n <- length(obs.data)
  x <- sum (obs.data)
  params <- betberPos (p1,p2,obs.data)
  a <- params[1]
  b <- params[2]

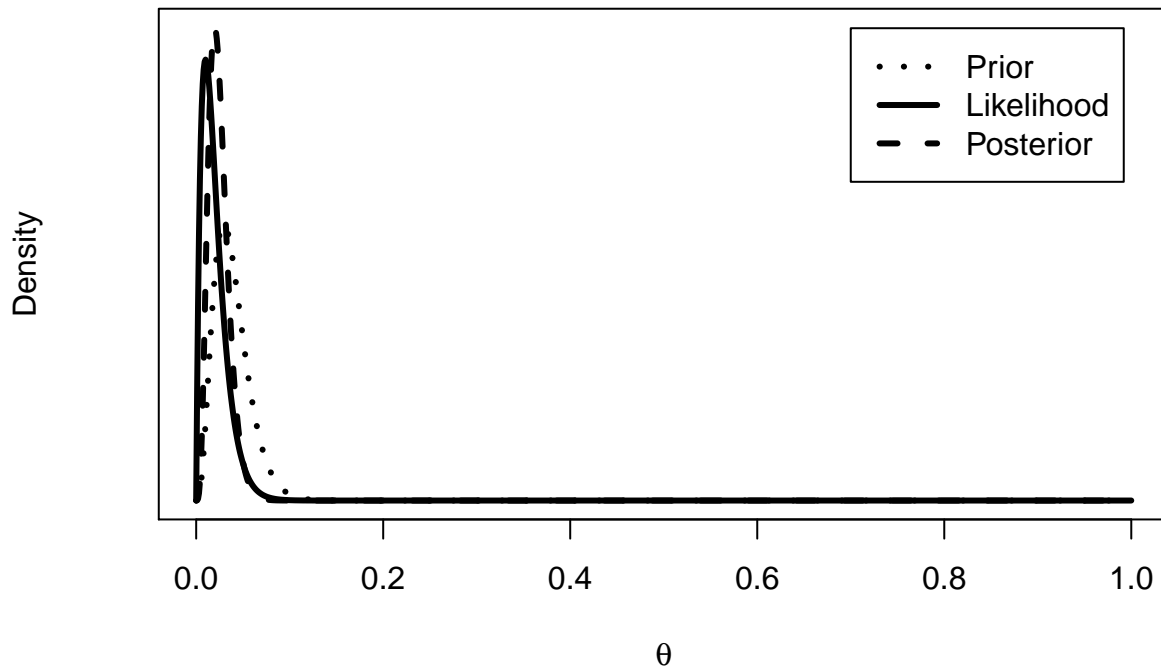
  #Define Lines
  like = dbeta(theta.sim, x+1, n-x+1)
  prior = dbeta(theta.sim, a, b)
  post = dbeta(theta.sim, x+a, n-x+b)

  #Plot Lines
  plot(theta.sim,post,type="l",ylab="Density", yaxt = 'n',lty=2,lwd=3,
  xlab = expression(theta))
  lines(theta.sim,like,lty=1,lwd=3)
  lines(theta.sim,prior,lty=3,lwd=3)
  legend(0.7,40,c("Prior","Likelihood","Posterior"),
  lty=c(3,1,2),lwd=c(3,3,3))
} # yaxt from: https://stackoverflow.com/questions/10393076/suppress-ticks-in-plot-in-r

distPlot(1,1)
```



```
distPlot(3,1)
```



2. {The Exponential-Gamma Model}

We write $X \sim \text{Exp}(\theta)$ to indicate that X has the Exponential distribution, that is, its p.d.f. is

$$p(x|\theta) = \text{Exp}(x|\theta) = \theta \exp(-\theta x) \mathbf{1}(x > 0).$$

The Exponential distribution has some special properties that make it a good model for certain applications. It has been used to model the time between events (such as neuron spikes, website hits, neutrinos captured in a detector), extreme values such as maximum daily rainfall over a period of one year, or the amount of time until a product fails (lightbulbs are a standard example).

Suppose you have data x_1, \dots, x_n which you are modeling as i.i.d. observations from an Exponential distribution, and suppose that your prior is $\theta \sim \text{Gamma}(a, b)$, that is,

$$p(\theta) = \text{Gamma}(\theta|a, b) = \frac{b^a}{\Gamma(a)} \theta^{a-1} \exp(-b\theta) \mathbb{1}(\theta > 0).$$

- a. Derive the formula for the posterior density, $p(\theta|x_{1:n})$. Give the form of the posterior in terms of one of the most common distributions (Bernoulli, Beta, Exponential, or Gamma).

By Baye's Theorem:

$$\begin{aligned} p(\theta|x_{1:n}) &= \frac{p(x_{1:n}|\theta)p(\theta)}{p(x_{1:n})} \propto p(x_{1:n}|\theta)p(\theta) \\ p(x_{1:n}|\theta) &= \prod_{i=1}^n \theta e^{-\theta x_i} \\ p(x_{1:n}|\theta) &= \theta^n e^{-\theta \sum x_i} \\ p(\theta|x_{1:n}) &= \theta^n e^{-\theta \sum x_i} \frac{b^a}{\Gamma(a)} \theta^{a-1} e^{-b\theta} \\ \theta^n e^{-\theta \sum x_i} \frac{b^a}{\Gamma(a)} \theta^{a-1} e^{-b\theta} &= \frac{b^a}{\Gamma(a)} \theta^{n+a-1} e^{-\theta \sum x_i - b\theta} \\ \frac{b^a}{\Gamma(a)} \theta^{n+a-1} e^{-\theta \sum x_i - b\theta} &= \frac{b^a}{\Gamma(a)} \theta^{n+a-1} e^{-\theta \sum x_i + b(-\theta)} \\ p(\theta|x_{1:n}) &= \text{Gamma}(a + n, \sum x_i + b) \end{aligned}$$

- b. Why is the posterior distribution a *proper* density or probability distribution function?

The posterior distribution is a Gamma distribution, which is a proper probability distribution function with an area under the curve of 1.

- c. Now, suppose you are measuring the number of seconds between lightning strikes during a storm, your prior is $\text{Gamma}(0.1, 1.0)$, and your data is

$$(x_1, \dots, x_8) = (20.9, 69.7, 3.6, 21.8, 21.4, 0.4, 6.7, 10.0).$$

Plot the prior and posterior p.d.f.s. (Be sure to make your plots on a scale that allows you to clearly see the important features.)

```
obs.data <- c(20.9, 69.7, 3.6, 21.8, 21.4, 0.4, 6.7, 10.0) #Define data

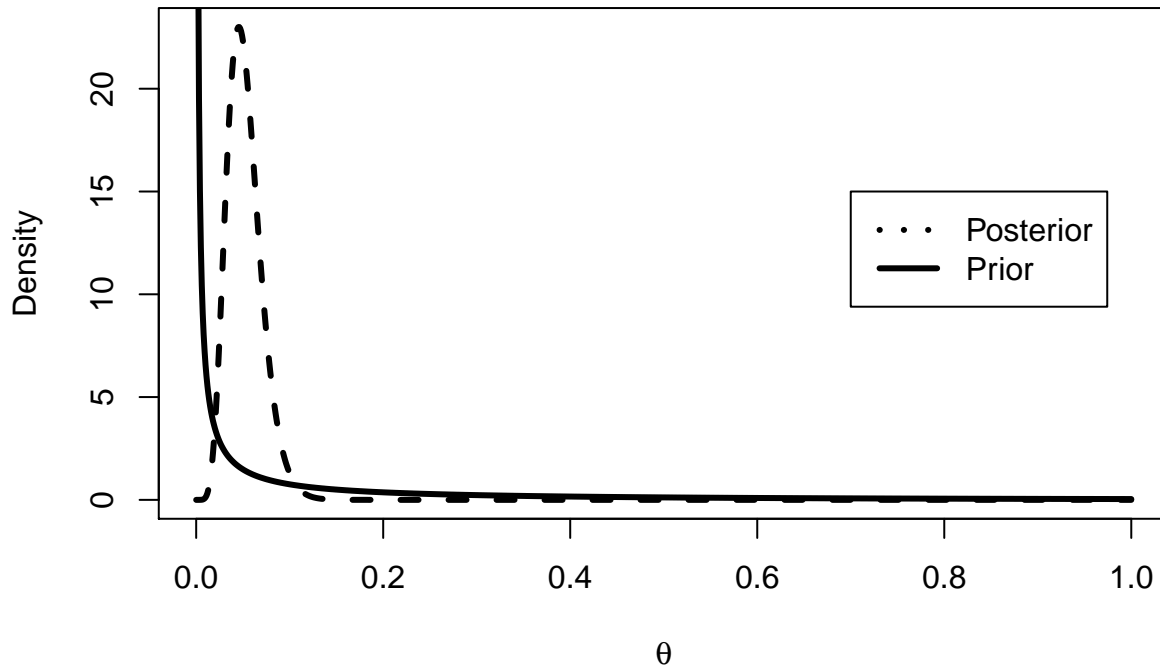
#Define params
x <- sum (obs.data)
n <- length (obs.data)

a <- 0.1
b <- 1.0

#Define Lines
prior = dgamma(theta.sim, a, b)
post = dgamma(theta.sim, a+n, x+b)

#Plot Lines
plot(theta.sim, post, type="l", ylab="Density", lty=2, lwd=3,
xlab = expression(theta))
```

```
lines(theta.sim,prior,lty=1,lwd=3)
legend(0.7,15,c("Posterior", "Prior"),
lty=c(3,1,2),lwd=c(3,3,3))
```



d. Give a specific example of an application where an Exponential model would be reasonable. Give an example where an Exponential model would NOT be appropriate, and explain why.

An Exponential distribution would be reasonable to model the time between car accidents on a stretch of road, as a recent car accident would not change the likelihood of car accidents after it (assuming pile-ups are counted as one accident).

It would not be reasonable to model the longevity of an automobile, because its parts would wear down over time and increase the likelihood of failure as time goes on.

3. {Priors, Posteriors, Predictive Distributions (Hoff, 3.9)}

An unknown quantity

$$Y \mid \theta$$

has a Galenshore(a, θ) distribution if its density is given by

$$p(y \mid \theta) = \frac{2}{\Gamma(a)} \theta^{2a} y^{2a-1} e^{-\theta^2 y^2}$$

for $y > 0, \theta > 0, a > 0$. Assume for now that a is known and θ is unknown and a random variable. For this density $E[Y] = \frac{\Gamma(a+1/2)}{\theta \Gamma(a)}$ and $E[Y^2] = \frac{a}{\theta^2}$.

- Identify a class of conjugate prior densities for θ . Assume the prior parameters are c and d . That is, state the distribution that θ should have with parameters c, d such that the resulting posterior is conjugate. Plot a few members of this class of densities.

If the prior is a *Galenshore* distribution, then the posterior distribution should also be a *Galenshore* distribution, and therefore conjugate (since the likelihood is a *Galenshore* distribution as well).

```

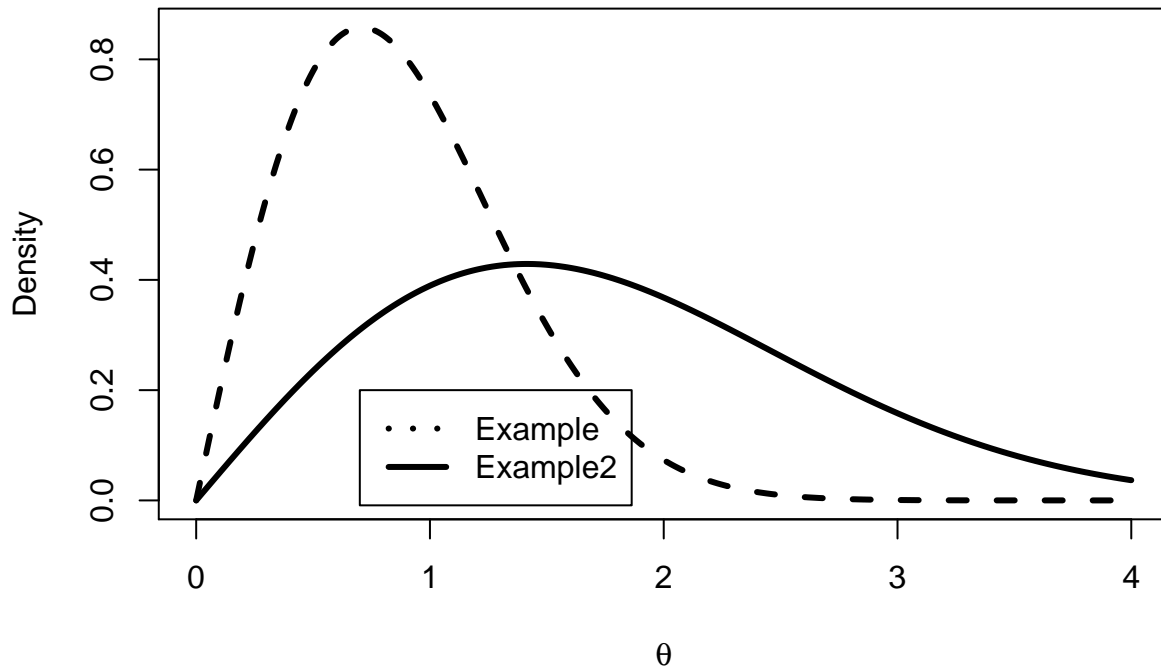
#Redefine theta to cover the range of my sample Galenshore functions
theta.sim <- seq(from = 0, to = 4, length.out = 1000)

#Function that gets y-values for galenshore
galenshore <- function(th,c,d){
  y <- (2/gamma(c))*((d^(2*c))*(th^((2*c)-1))*exp((-d^2))* (th^2)))
  return(y)
}

#Define Lines
example = galenshore(theta.sim, 1, 1)
example2 = galenshore(theta.sim, 1, 0.5)

#Plot Lines
plot(theta.sim,example,type="l",ylab="Density",lty=2,lwd=3,
      xlab = expression(theta))
lines(theta.sim,example2,lty=1,lwd=3)
legend(0.7,0.2,c("Example", "Example2"),
      lty=c(3,1,2),lwd=c(3,3,3))

```



- b. Let $Y_1, \dots, Y_n \stackrel{iid}{\sim} \text{Galenshore}(a, \theta)$. Find the posterior distribution of $\theta \mid y_{1:n}$ using a prior from your conjugate class.

$$\begin{aligned}
 p(\theta \mid y_{1:n}) &\propto p(y_{1:n} \mid \theta)p(\theta) \\
 p(y_{1:n} \mid \theta)p(\theta) &= \left(\frac{2}{\Gamma(a)}\right)^n \theta^{2an} \left(\prod_{i=1}^n y_i^{2a-1}\right) e^{-\theta^2 \sum_{i=1}^n y_i^2} * \frac{2}{\Gamma(c)} d^{2c} \theta^{2c-1} e^{-d^2 \theta^2} \\
 &\propto \theta^{2an} e^{-\theta^2 \sum_{i=1}^n y_i^2} * \theta^{2c-1} e^{-d^2 \theta^2} \\
 p(\theta \mid y_{1:n}) &= \theta^{2an+2c-1} e^{-d^2 \theta^2 - \theta^2 \sum_{i=1}^n y_i^2} \\
 p(\theta \mid y_{1:n}) &= \theta^{2(an+c)-1} e^{-\theta(d^2 + \sum_{i=1}^n y_i^2)}
 \end{aligned}$$

c. Show that

$$\frac{p(\theta_a | y_{1:n})}{p(\theta_b | y_{1:n})} = \left(\frac{\theta_a}{\theta_b} \right)^{2(an+c)-1} e^{(\theta_b^2 - \theta_a^2)(d^2 + \sum y_i^2)},$$

where $\theta_a, \theta_b \sim \text{Galenshore}(c, d)$. Identify a sufficient statistic.

$$\begin{aligned} \frac{p(\theta_a | y_{1:n})}{p(\theta_b | y_{1:n})} &= \frac{\theta_a^{2(an+c)-1} e^{-\theta_a(d^2 + \sum_{i=1}^n y_i^2)}}{\theta_b^{2(an+c)-1} e^{-\theta_b(d^2 + \sum_{i=1}^n y_i^2)}} \\ \frac{p(\theta_a | y_{1:n})}{p(\theta_b | y_{1:n})} &= \left(\frac{\theta_a}{\theta_b} \right)^{2(an+c-1)} e^{(\theta_b^2 - \theta_a^2)(d^2 + \sum_{i=1}^n y_i^2)} \end{aligned}$$

$$\sum_{i=1}^n y_i^2$$

is a sufficient statistic for this problem because it is the only information from the sample that is necessary to calculate the answer.

d. Determine $E[\theta | y_{1:n}]$.

$$p(\theta | y_{1:n}) = \theta^{2(an+c)-1} e^{-\theta(d^2 + \sum_{i=1}^n y_i^2)}$$

This can be interpreted as $\text{Gamma}(an + c - \frac{1}{2}, d + \sum y_i)$

$$\text{Therefore } E[\theta | y_{1:n}] = \frac{\Gamma(an+c)}{d + \sum y_i \Gamma(an+c-\frac{1}{2})}$$

e. Show that the form of the posterior predictive density:

$$p(y_{n+1} | y_{1:n}) = \frac{2y_{n+1}^{2a-1} \Gamma(an+a+c)}{\Gamma(a) \Gamma(an+c)} \frac{(d^2 + \sum y_i^2)^{an+c}}{(d^2 + \sum y_i^2 + y_{n+1}^2)^{(an+a+c)}}.$$

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

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

$$p(y | \theta) = \int_{\theta} \frac{2}{\Gamma(a)} \theta^{2a} y_{n+1}^{2a-1} e^{-\theta^2 y_{n+1}^2} * \frac{2}{\Gamma(an+c)} \left(\sqrt{d^2 + \sum y_i^2} \right)^{2(an+c)} \theta^{2(an+c)-1} e^{-\left(\sqrt{d^2 + \sum y_i^2} \right)^2 \theta^2} d\theta$$

$$\begin{aligned} &= \frac{2y_{n+1}^{2a-1} 2 \left(\sqrt{d^2 + \sum y_i^2} \right)^{2(an+c)}}{\Gamma(a) \Gamma(an+c)} * \int_{\theta} \theta^{2(an+a+c)-1} e^{-\theta^2 [y_{n+1}^2 + d^2 + \sum y_i^2]} d\theta * \frac{(d^2 + \sum y_i^2 + y_{n+1}^2)^{(an+a+c)}}{(d^2 + \sum y_i^2 + y_{n+1}^2)^{(an+a+c)}} \frac{2}{\Gamma(an+a+c)} \frac{\Gamma(-)}{2} \\ &= \frac{2y_{n+1}^{a-1} 2(d^2 + \sum y_i^2)^{(an+c)}}{\Gamma(a) \Gamma(an+c)} * \int \frac{2}{\Gamma(an+a+c)} (d^2 + \sum y_i^2 + y_{n+1}^2)^{(an+a+c)} d\theta * \left[\frac{\Gamma(an+a+c)}{2} * \frac{1}{d^2 + \sum y_i^2 + y_{n+1}^2} \right] \\ &= \frac{2y_{n+1}^{a-1} (d^2 + \sum y_i^2)^{an+c} \Gamma(an+a+c)}{\Gamma(a) \Gamma(an+c) (d^2 + \sum y_i^2 + y_{n+1}^2)} \end{aligned}$$