HW2

STA-360/602, Spring 2018

Frank

1/16/2019

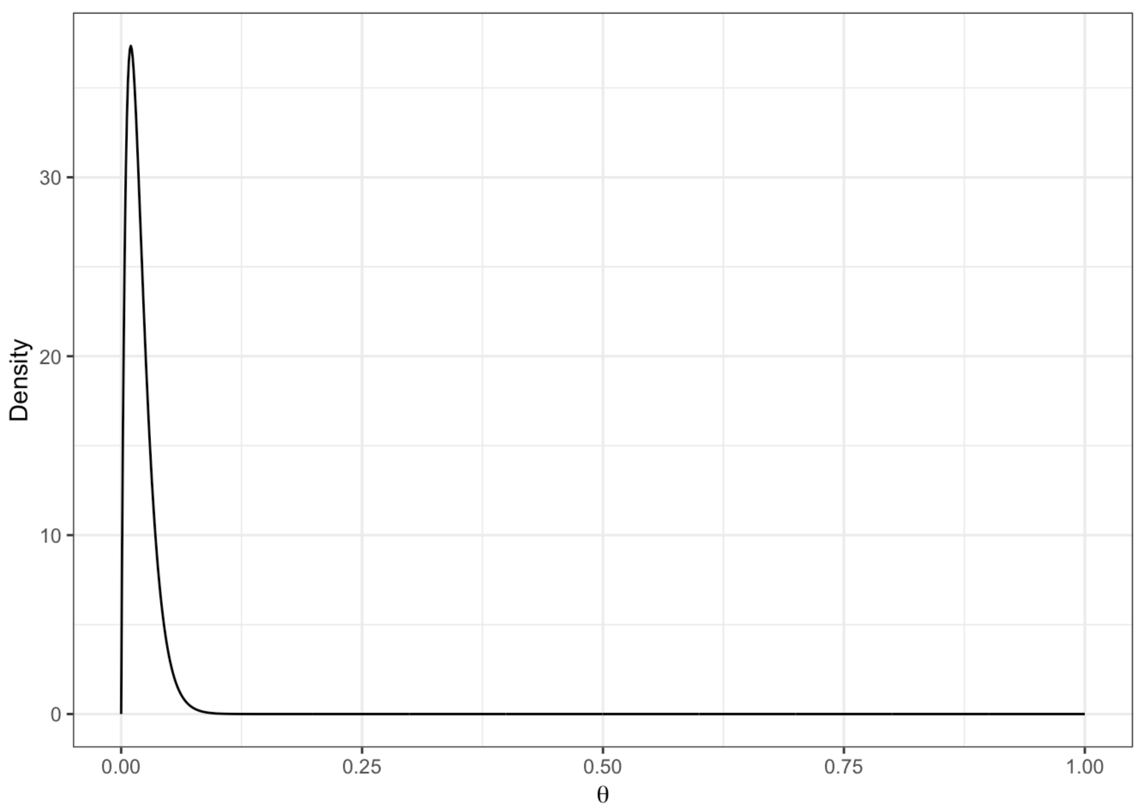
1. Lab component

(a) 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.

```
# Generating simulated data
set.seed(123)
obs.data <- rbinom(n = 100, size = 1, prob = 0.01)
head(obs.data)
```

```
## [1] 0 0 0 0 0 0
```

```
# Create and plot likelihood function
n = length(obs.data)
x = sum(obs.data)
th = seq(0, 1, length = 1000)

like = dbeta(th, x+1, n-x+1)
ggplot()+geom_line(aes(x = th, y = like))+labs(x = expression(theta), y = "Density")+ theme_bw()
```



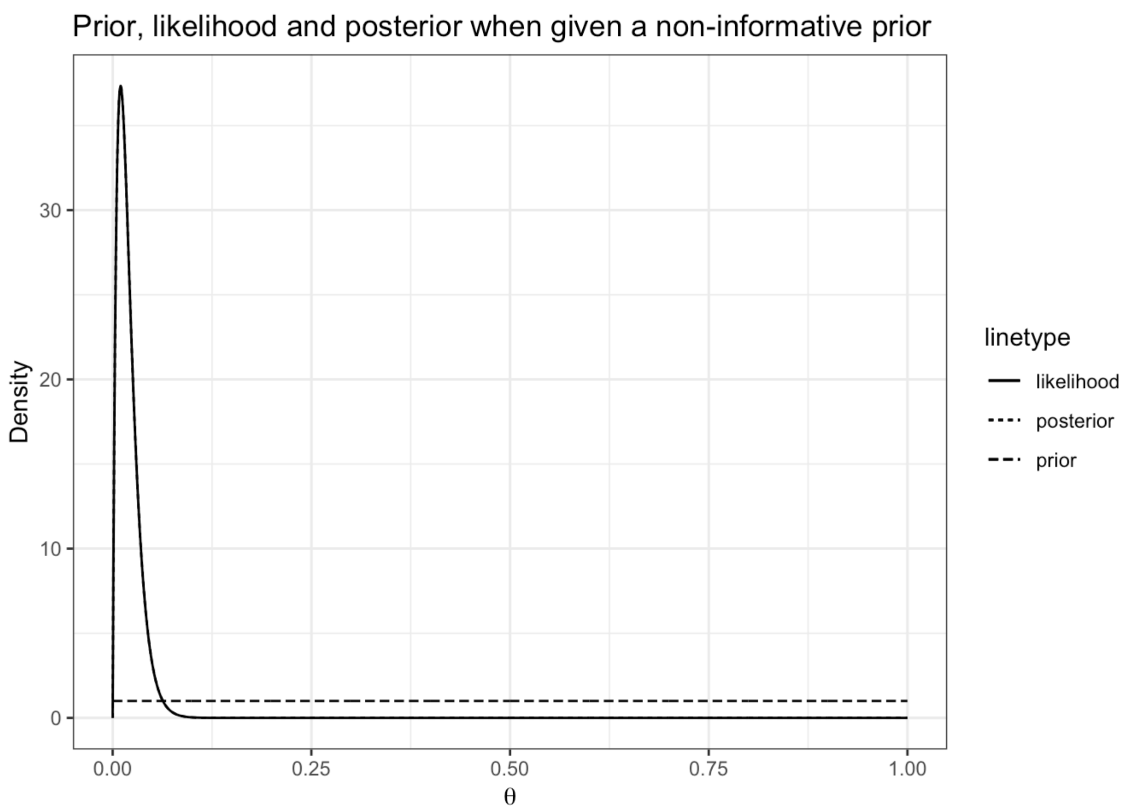
(b) Task 4. Write a function that takes as its inputs prior parameters a and b for the Beta-Bernoulli model and the observed data, and produces the posterior parameters you need for the model. Generate the posterior parameters for a non-informative prior i.e. $(a, b) = (1, 1)$ and for an informative case

$(a, b) = (3, 1)$

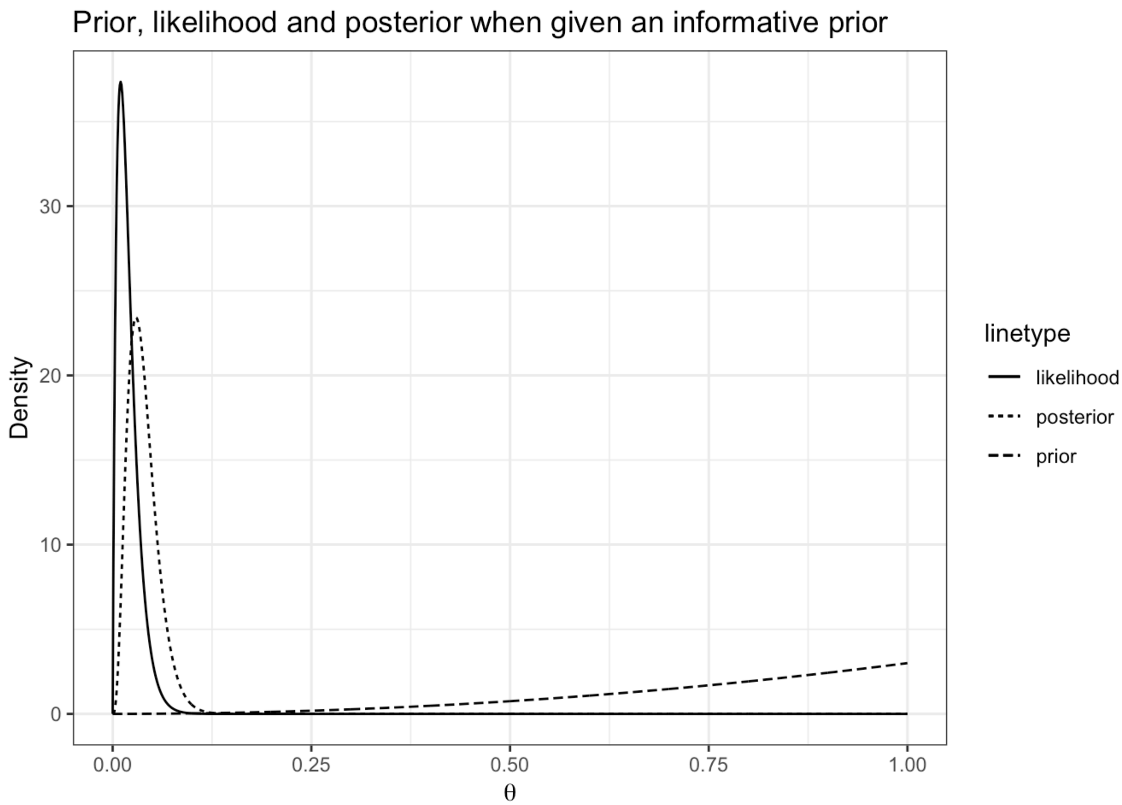
```
a1 = 1
b1 = 1
a2 = 3
b2 = 1

prior1 = dbeta(th, a1, b1)
post1 = dbeta(th, a1+x, b1+n-x)

prior2 = dbeta(th, a2, b2)
post2 = dbeta(th, a2+x, b2+n-x)
```

(c) Task 5. Create two plots, one for the informative and one for the non-informative case to show the posterior distribution and superimpose the prior distribu- tions 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.

```
# Plot for non-informative case
ggplot() +
  geom_line(aes(x = th, y = post1, linetype = 'posterior'))+
  geom_line(aes(x = th, y = prior1, linetype = 'prior'))+
  geom_line(aes(x = th, y = like, linetype = 'likelihood'))+
  labs(x = expression(theta), y = "Density", title = "Prior, likelihood and posterior when given a non-infor
mative prior") + theme_bw()
```




```
# Plot for informative case
ggplot() +
  geom_line(aes(x = th, y = post2, linetype = 'posterior'))+
  geom_line(aes(x = th, y = prior2, linetype = 'prior'))+
  geom_line(aes(x = th, y = like, linetype = 'likelihood'))+
  labs(x = expression(theta), y = "Density", title = "Prior, likelihood and posterior when given an informat
ive prior") + theme_bw()
```



2. The *Exponential-Gamma Model* We write $X \sim Exp(\theta)$ to indicate that X has the Exponential distribution, that is, its p.d.f. is

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

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 $x_1, \ldots, x_n$ which you are modeling as i.i.d. observations from an Exponential distribution, and suppose that your prior is $\theta \sim Gamma(a, b)$, that is,

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

(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).

$$\begin{aligned}
p(\theta|x_{1:n}) &\propto p(x_{1:n}|\theta)p(\theta) \\
&\propto \prod_n [\theta \exp(-\theta x)] \frac{b^a}{\Gamma(a)} \theta^{a-1} \exp(-b\theta) \\
&\propto \theta^n \exp(-n\bar{x}\theta)\theta^{a-1} \exp(-b\theta) \\
&\propto \theta^{n+a-1} \exp[-(n\bar{x} + b)\theta] \\
&\propto Gamma(\theta|n + a, n\bar{x} + b)
\end{aligned}$$

(b) Why is the posterior distribution a *proper* density or probability distribution function?

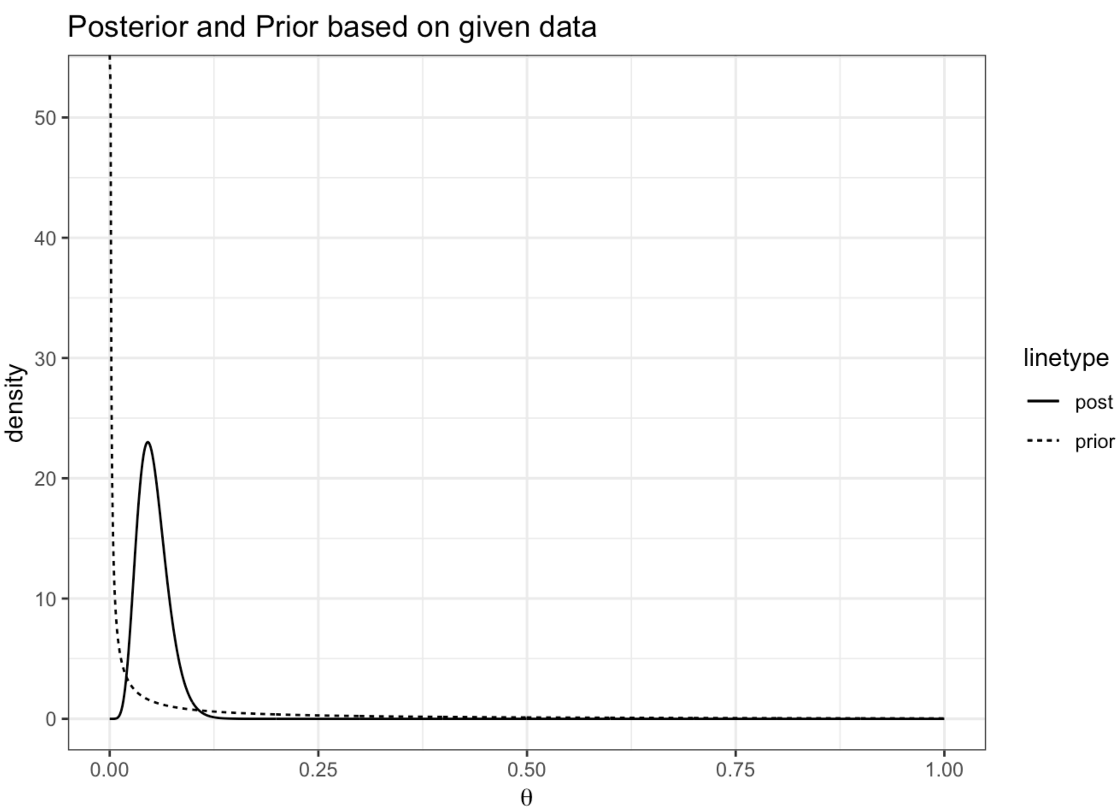
Because the prior $p(\theta) = Gamma(\theta|a, b)$ is a conjugate prior for the given likelihood function. That means the posterior is in the same probability distribution family as the prior probability distribution $p(\theta)$ is.

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

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

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.)

```
X = c(20.9, 69.7, 3.6, 21.8, 21.4, 0.4, 6.7, 10.0)
th = seq(0,1,length = 1000)
x <- sum(X)
n <- length(X)
a = 0.1
b = 1.0

# Generate the prior and post function according to conclusions above
prior <- dgamma(th, a, b)
post <- dgamma(th, a + n, b + x)

# Plotting
ggplot()+geom_line(aes(x = th, y = prior, linetype = 'prior'))+
  geom_line(aes(x = th, y = post, linetype = 'post'))+
  labs(x = expression(theta), y = 'density', title = 'Posterior and Prior based on given data') + theme_bw()
```



(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.

Reasonable: The time that some radioactive matter decays.

Unreasonable: The amount of flying bugs in different hours of day. Because most bugs are only active during night time, which is not constant throughout the day.

3. *Priors, Posteriors, Predictive Distributions (Hoff, 3.9)*.

An unknown quantity Y has a Galenshore(a, θ) distribution if its density is given by

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

for $y > 0, \theta > 0, a > 0$. Assume for now that a is known. For this density,

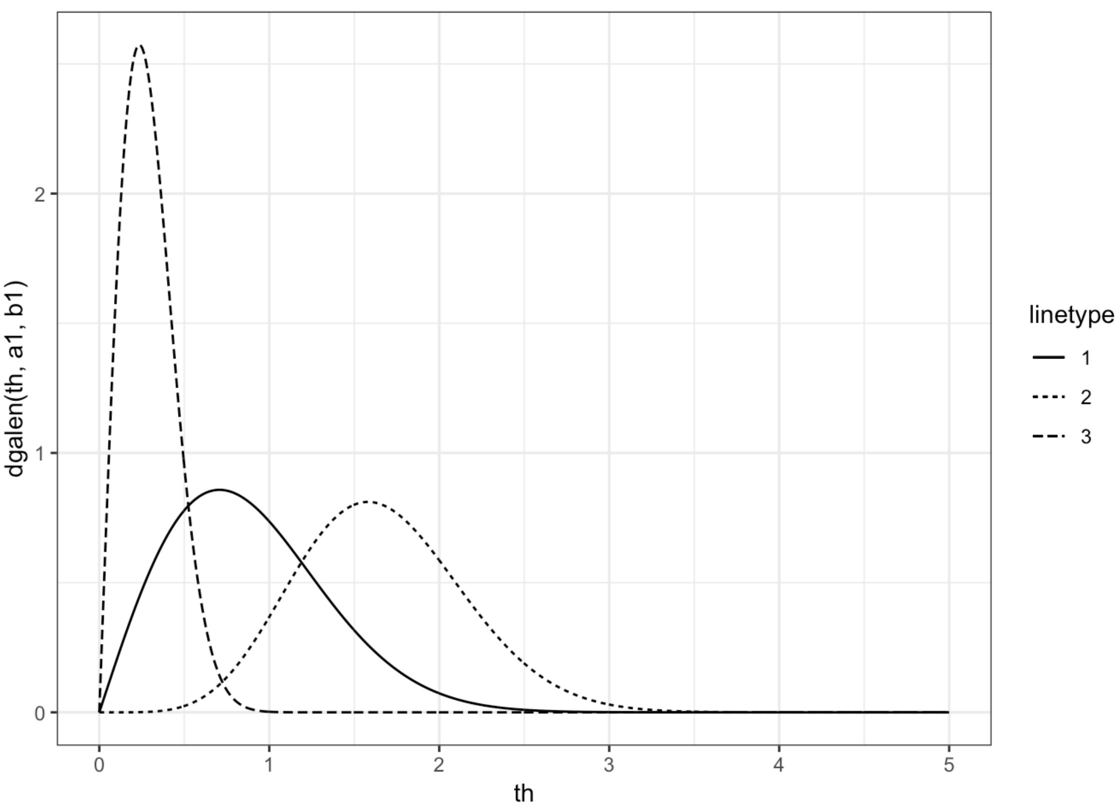
$$E[Y] = \frac{\Gamma(a+1/2)}{\theta\Gamma(a)}$$

and

$$E[Y^2] = \frac{a}{\theta^2}.$$

(a) Identify a class of conjugate prior densities for θ. Plot a few members of this class of densities.

```
# Define dgalen function
dgalen <- function(th,a,b){
  return(2/gamma(a)*b^(2*a)*th^(2*a-1)*exp(-b^2*th^2))
}
# Creating a set of (a, b) and theta for plotting
a1 = 1
b1 = 1
a2 = 3
b2 = 1
a3 = 1
b3 = 3
th = seq(0, 5, length = 1000)

# Plotting
ggplot()+geom_line(aes(x=th, y=dgalen(th,a1,b1),linetype='1'))+
  geom_line(aes(x=th, y=dgalen(th,a2,b2),linetype='2'))+
  geom_line(aes(x=th, y=dgalen(th,a3,b3),linetype='3'))+
  theme_bw()
```



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

Suppose the prior, θ, is $\theta \sim$ Galenshore(k, b),

Given that Y has a Galenshore(a, θ) distribution for y1,…,yn,

then the posterior is:

$$
\begin{aligned}
p(\theta|y_{1:n}) &\propto p(\theta)p(y_{1:n}|\theta)\\
&\propto \theta^{2k-1}\exp(-b^2\theta^2)\prod_n[\theta^{2a}yi^{2a-1}\exp(-\theta^2 yi^2)]\\
&\propto \theta^{2k-1}\exp(-b^2\theta^2)\theta^{2an}\exp(-\theta^2\sum_n yi^2)\\
&\propto \theta^{2k+2an-1}\exp[-\theta^2(b^2+\sum_n yi^2)]\\
&\propto Galenshore(\alpha=k+an,\beta=\sqrt{b^2+\sum_n yi^2})
\end{aligned}
$$

(c) Write down

$$\frac{p(\theta_a \mid y_{1:n})}{p(\theta_b \mid y_{1:n})}$$

and simplify. Identify a sufficient statistic.

Given the result from (b), $p(\theta|y_{1:n}) \propto Galenshore(\alpha = k + an, \beta = \sqrt{b^2 + \sum_n yi^2})$:

$$\frac{p(\theta_a \mid y_{1:n})}{p(\theta_b \mid y_{1:n})} = \frac{\theta_a^{2\alpha-1}exp(-\beta^2\theta_a^2)}{\theta_b^{2\alpha-1}exp(-\beta^2\theta_b^2)} = \left(\frac{\theta_a}{\theta_b}\right)^{2\alpha-1} exp[-\beta^2(\theta_a^2-\theta_b^2)]$$

Sufficient statistic: $\sum_n yi^2$

(d) Determine $E[\theta \mid y_{1:n}]$.

Since this is a conjugate family, and since given $Y \sim Galenshore(a, \theta)$,

$$
\begin{gathered}
E[Y] = \frac{\Gamma(a+1/2)}{\theta\Gamma(a)}\\
E[\theta \mid y_{1:n}] = \frac{\Gamma(\alpha+1/2)}{\beta\Gamma(\alpha)}\\
given: \alpha = k + an, \beta = \sqrt{b^2+\sum_n yi^2},\\
where: \theta \sim Galenshore(k, b)
\end{gathered}
$$

(e) Determine the form of the posterior predictive density $p(y_{n+1} \mid y_{1:n})$

$$
\begin{aligned}
p(y_{n+1} \mid y_{1:n}) &= \int p(y_{n+1}|\theta)p(\theta|y_{1:n})d\theta\\
&= \int \frac{2}{\Gamma(a)}\theta^{2a}y_{n+1}^{2a-1}e^{-\theta^2 y_{n+1}^2}\frac{2}{\Gamma(\alpha)}\beta^{2\alpha}\theta^{2\alpha-1}e^{-\beta^2\theta^2}d\theta\\
&= \frac{4}{\Gamma(a)\Gamma(\alpha)}y_{n+1}^{2a-1}\beta^{2\alpha}\int\theta^{2(a+\alpha)-1}e^{-\theta^2(y_{n+1}^2+\beta^2)}d\theta\\
&= \frac{4}{\Gamma(a)\Gamma(\alpha)}y_{n+1}^{2a-1}\beta^{2\alpha}\int\frac{\Gamma(a+\alpha-\frac{1}{2})}{2}(y_{n+1}^2+\beta^2)^{-(a+\alpha-\frac{1}{2})}\frac{2}{\Gamma(a+\alpha-\frac{1}{2})}(y_{n+1}^2+\beta^2)^{a+\alpha-\frac{1}{2}}\theta^{2(a+\alpha-\frac{1}{2})-1}e^{-\theta^2(y_{n+1}^2+\beta^2)}\theta d\theta\\
&= \frac{2\Gamma(a+\alpha-\frac{1}{2})}{\Gamma(a)\Gamma(\alpha)}y_{n+1}^{2a-1}\beta^{2\alpha}(y_{n+1}^2+\beta^2)^{-(a+\alpha-\frac{1}{2})}\int Galenshore(\theta|a+\alpha-\frac{1}{2},\sqrt{y_{n+1}^2+\beta^2})\theta d\theta\\
&\propto y_{n+1}^{2a-1}(y_{n+1}^2+\beta^2)^{-(a+\alpha-\frac{1}{2})}\frac{\Gamma(a+\alpha)}{\sqrt{y_{n+1}^2+\beta^2}\Gamma(a+\alpha-\frac{1}{2})}\\
&\propto y_{n+1}^{2a-1}(y_{n+1}^2+\beta^2)^{-(a+\alpha)}
\end{aligned}
$$

This is the simpliest form I can get, but I cannot recognize the family of the distribution.