Bayesian Cat I

124384 Luycer Bosire

2/6/2021

**Question 1: Single parameter poisson**

Consider the count of airline crashes per year over a 10-year period form 1976 to 1985:

y <- c(24,25,31,31,22,21,26,20,16,22)  
n<- length(y)  
n  
bar\_x<- mean(y)  
bar\_x

***a. Show that the Gamma distribution is the conjugate prior for a Poisson mean.***

Solution in the steps below;

1. write down the likelihood for .

The pdf and likelihood function of Poisson Distribution with parameter is given by;

1. Write down the prior density for .

The prior density Gamma with parameter

is given by;

Ignoring the constant term:

3. Multiply them together to obtain the posterior density, and notice that it has the same form as the gamma distribution.

Using Bayes Rule

That is

Therefore:

Thus we can notice that the posterior takes a Gamma form:

Where

***b. Show graphics of the resulting posterior distribution for choices of various gamma priors***

Confirming the code for different values of alpha and beta:

library(shiny)  
library(tidyverse)

## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --

## v ggplot2 3.3.3 v purrr 0.3.4  
## v tibble 3.0.5 v dplyr 1.0.3  
## v tidyr 1.1.2 v stringr 1.4.0  
## v readr 1.4.0 v forcats 0.5.1

## -- Conflicts ------------------------------------------ tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

# ui codes ----------------------------------------------------------------  
  
ui <- fluidPage(  
 sidebarLayout(  
 sidebarPanel(  
 sliderInput("n", "Play to increase sample size", min = 1,  
 max = 10, step = 1,value = 10, animate = T),  
 tags$hr(),  
 numericInput("beta", "Beta value", value = 5),  
 tags$hr(),  
 numericInput("alpha", "Alpha Value", value = 5), width = 2  
 ),  
   
 mainPanel(  
 h3("The likelihood and posterior distributions converge as the sample size n increases"),  
 br(),  
plotOutput("my\_plot", height = "500px")  
 )  
   
 )  
)  
  
# server code -------------------------------------------------------------  
  
server <- function(input, output, session){  
   
  
  
# define posterior function -----------------------------------------------  
y <- c(24,25,31,31,22,21,26,20,16,22)  
Lambda <- mean(y)  
post <- function(n = 10,alpha = 5, beta = 1/Lambda){  
  
lambda<-Lambda # 1data initialization  
  
LL <- dpois(y,lambda) # quartiles of a binomial  
  
# prior  
Prior <- dgamma(y,shape = alpha,scale = beta) #beta prior distribution  
   
# posterior  
alpha1 <- n\*lambda + alpha  
beta1 <- n+ beta  
  
Postr <- dgamma(y,shape=alpha1,scale = beta1)  
  
ggplot(data = NULL, aes(y, LL/max(LL), col = "Likelihood")) + geom\_line(size = 1.0)+  
 geom\_line(aes(y, Prior/max(Prior), col = "Prior"), size = 1.0) +   
 geom\_line(aes(y, Postr/max(Postr), col = "Posterior"), size = 1.0) +  
 labs(y = "Density", x= expression(y)) + theme\_minimal() +  
 scale\_colour\_manual("", values = c("Likelihood"="purple", "Prior"="red", "Posterior"="blue")) +  
 theme(legend.position = "top")  
  
}  
  
output$my\_plot <- renderPlot({  
 post(n=input$n,alpha = input$alpha, beta = input$beta)  
})  
   
   
}  
  
shinyApp(ui, server)

***c. Show a table of the prior, MLE and posterior estimates of the poison mean under different choices of the gamma priors in (a) above.***

1.When and

Table\_Estimates <- function(alpha,beta){  
   
y <- c(24,25,31,31,22,21,26,20,16,22)  
lambda <- mean(y)  
MLE <- dpois(y,lambda)  
Prior01 <- dgamma(y,shape = alpha,scale = beta)  
Postr01 <- MLE\*Prior01  
  
Estimates <- rbind(MLE,Prior01,Postr01)  
  
return(Estimates)  
  
}  
df <- data.frame(Table\_Estimates(3,6))  
df

## X1 X2 X3 X4 X5  
## MLE 0.081083527 0.077191517 0.0264639843 0.0264639843 0.079016501  
## Prior01 0.024420852 0.022430344 0.0126877560 0.0126877560 0.028638384  
## Postr01 0.001980129 0.001731432 0.0003357686 0.0003357686 0.002262905  
## X6 X7 X8 X9 X10  
## MLE 0.073040463 0.070659927 0.064447467 0.0233562465 0.079016501  
## Prior01 0.030826496 0.020536205 0.033031475 0.0411753785 0.028638384  
## Postr01 0.002251582 0.001451087 0.002128795 0.0009617023 0.002262905

**QUESTION TWO**

***a.Differentiate between Credible Intervals and the Highest Posterior Density (HPD) in Bayesian analysis.***

A credible interval is an interval within which an unobserved parameter value falls with a particular probability. It is an interval in the domain of a posterior probability distribution or a predictive distribution.On the other hand, a highest posterior density (interval) is basically the shortest interval on a posterior density for some given confidence level.

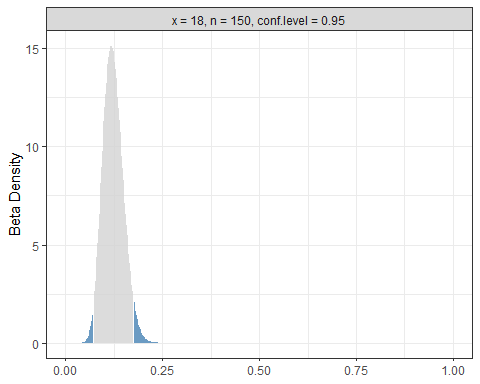
***b.Graphically demonstrate this for a disease prevalence given that out of 150 individuals 18 had the disease. Hint: assume Binomial distribution***

Highest Posterior Density (HPD)

library(binom)  
HPD <- binom.bayes(x=18,n=150,type = "highest",conf.level = 0.95,tol = 1e-9)  
print(HPD)

## method x n shape1 shape2 mean lower upper sig  
## 1 bayes 18 150 18.5 132.5 0.1225166 0.07246161 0.1754186 0.05

binom.bayes.densityplot(HPD)



Credible Interval

library(binom)  
Central <- binom.bayes(x=18,n=150,type = "central",conf.level = 0.95,tol = 1e-9)  
print(Central)

## method x n shape1 shape2 mean lower upper sig  
## 1 bayes 18 150 18.5 132.5 0.1225166 0.07534776 0.179135 0.05

binom.bayes.densityplot(Central)

