

Bayesian Econometrics: Homework 2

Jerin Varghese

February 2, 2019

Problem 1

There are three boxes (A, B, and C) and one of the boxes has a gift certificate for a brand new iPhone XSMAX. Gheorghe, the winner of the lottery, is asked to pick which box holds the certificate. If he picks the right box, he will get to keep the prize, a new iPhone XSMAX. When Gheorghe picks a box, Partha, the host, opens an empty box and asks Gheorghe whether he wants to change his mind and pick the other box. Suppose Gheorghe picks box A. Partha shows box B empty and asks Gheorghe if he wants to change his mind and switch to box C. Use Bayes' Theorem to analyze whether Gheorghe should switch to box C.

Answer

Gheorghe chose box A in this situation. That is the first conditional we take in mind as we move forward. We assume that Partha would never open the box that has the iPhone in it. So, let's first define some events:

$\text{Box}[(A,B,C)] = \text{Partha opens box A/B/C}$

$\text{Phone}[(A,B,C)] = \text{iPhone is in box A/B/C}$

Let's now write out the probabilities we know:

- $P(\text{Box}[A]) = 0$.
 - Partha cannot open box A since that is the box Gheorghe selected.
- $P(\text{Box}[B]) = P(\text{Box}[C]) = 1/2$.
 - Partha could open either box B or C.
- $P(\text{Phone}[A]) = P(\text{Phone}[B]) = P(\text{Phone}[C]) = 1/3$
 - Chances of the iPhone being in any of the boxes is equal.
- $P(\text{Box}[B] \mid \text{Phone}[A]) = 1/2$
 - If the iPhone is in box A, the box Gheorghe selected, Partha will open either box B or C.
- $P(\text{Box}[B] \mid \text{Phone}[B]) = 0$
 - If the iPhone is in box B, Partha would never open box B.
- $P(\text{Box}[B] \mid \text{Phone}[C]) = 1$
 - If the iPhone is in box C, Partha has to open box B.

We're interested in $P(\text{Phone}[C] \mid \text{Box}[B])$. What is the probability that the iPhone is in box C, meaning that Gheorghe should switch, given that Partha opened box B? Applying Bayes' Theorem then...

$$P(\text{Phone}[C] \mid \text{Box}[B]) = (P(\text{Box}[B] \mid \text{Phone}[C]) * P(\text{Phone}[C])) / P(\text{Box}[B]) = (1 * 1/3) / (1/2) = 2/3$$

Similarly,

$$P(\text{Phone}[A] \mid \text{Box}[B]) = (P(\text{Box}[B] \mid \text{Phone}[A]) * P(\text{Phone}[A])) / P(\text{Box}[B]) = (1/2 * 1/3) / (1/2) = 1/3$$

Thus, we conclude that Gheorghe should switch to box C, since the probability that the iPhone is in box C is 2/3, while the probability that the iPhone is in box A is only 1/3.

Problem 2-1

Consider the following normal regression model:

$$y_i = \beta x_i + \epsilon_i \text{ where } \epsilon_i \text{ i.i.d. } N(0, \sigma^2)$$

Do the following using R:

- a) Generate 100 observations of x_i from the i.i.d. uniform distribution on the interval (0,20).

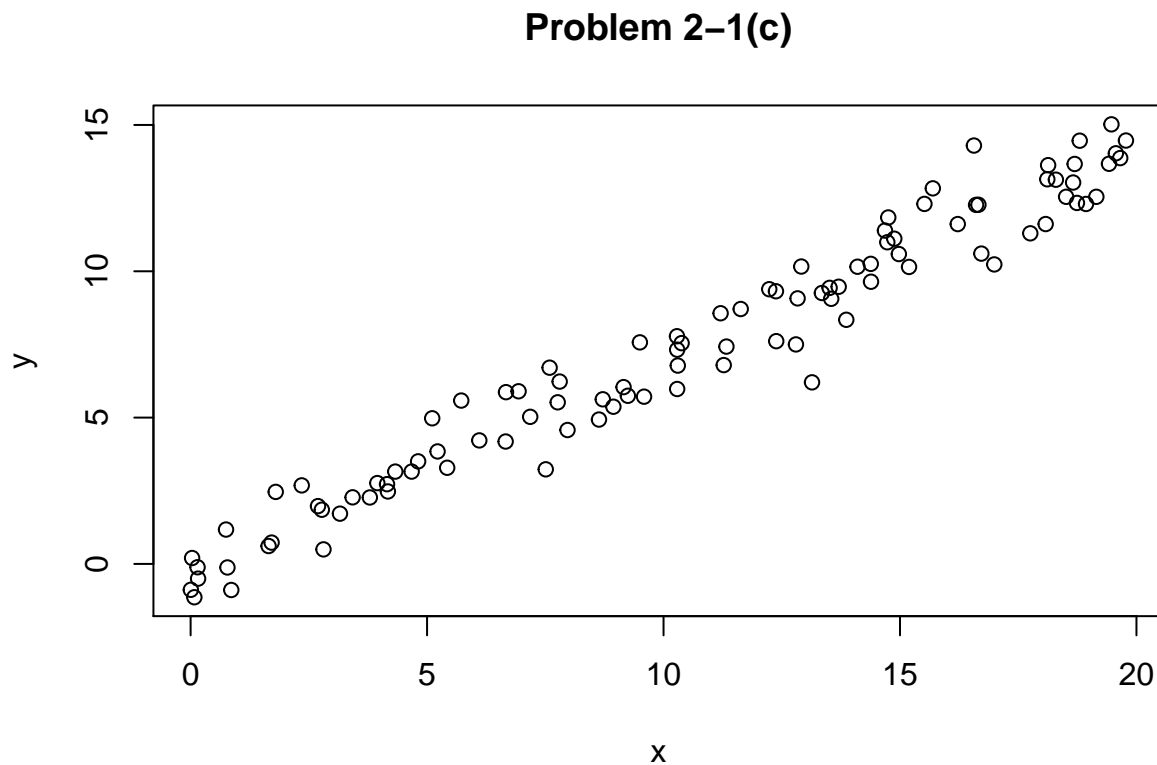
```
set.seed(42)
x <- runif(100, min=0, max=20)
```

- b) Then, generate 100 observations of y_i using the (x_i) 's drawn from a) with $\beta = 0.7$ and drawing 100 observations of ϵ_i with $\sigma = 1$.

```
beta <- 0.7
e <- rnorm(100, mean=0, sd=1)
y <- beta*x + e
```

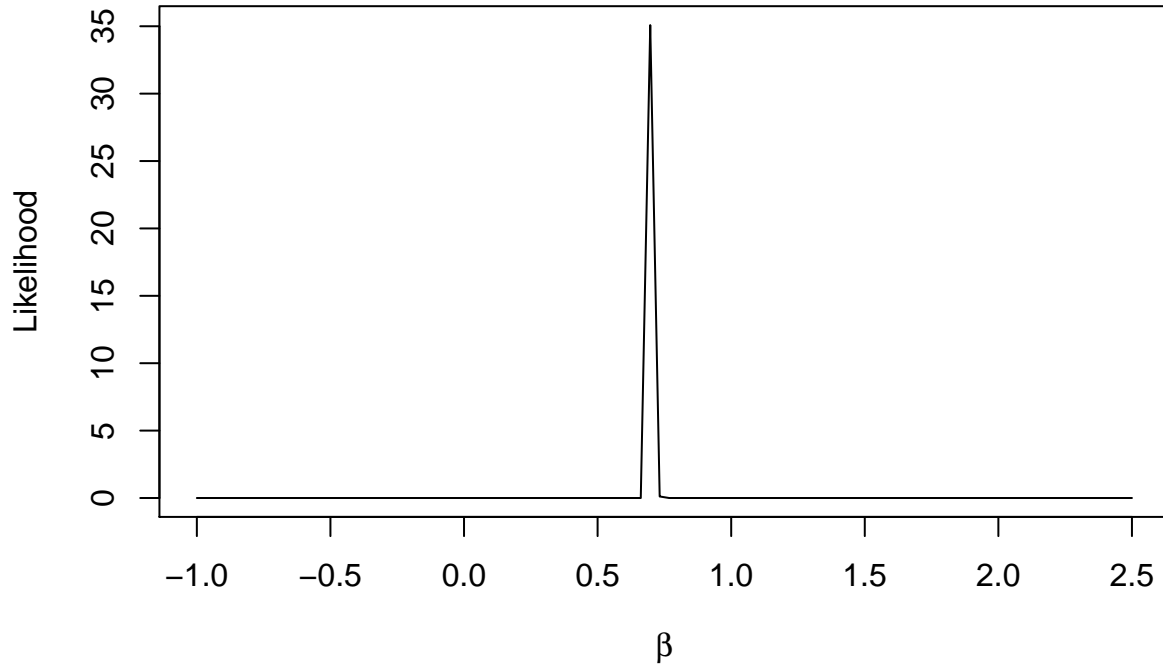
- c) Plot the data. And Plot the likelihood on the interval (-1, 2.5) for β

```
#plot data
plot(x, y, main="Problem 2-1(c)")
```



```
# plot likelihood
betahat <- sum(x*y)/sum(x*x)
betaval <- seq(-1,2.5,length=100)
invisible(plot(betaval,dnorm(betaval,betahat,1/sqrt(sum(x*x))),type="l", xlab="", ylab="")+
  title(main="Problem 2-1(c)", xlab=expression(beta), ylab="Likelihood"))
```

Problem 2-1(c)



Problem 2-2

a) Show that the likelihood for n independent realizations corresponding to n different y_i 's is given by:

$$L(\beta; y, x) = \exp\{-\beta \sum_{i=1}^n x_i\} \exp\{-\sum_{i=1}^n y_i e^{-\beta x_i}\}$$

Answer

$$L(\beta : y, x) = \prod_{i=1}^n \frac{1}{\mu} e^{-\frac{y}{\mu}}$$

Plugging in $\mu = e^{\beta x} \dots$

$$L(\beta : y, x) = \prod_{i=1}^n e^{-\beta x} e^{-\frac{y}{e^{\beta x}}}$$

$$L(\beta : y, x) = e^{-\beta \sum_{i=1}^n x_i} e^{-\sum_{i=1}^n y_i e^{-\beta x_i}}$$

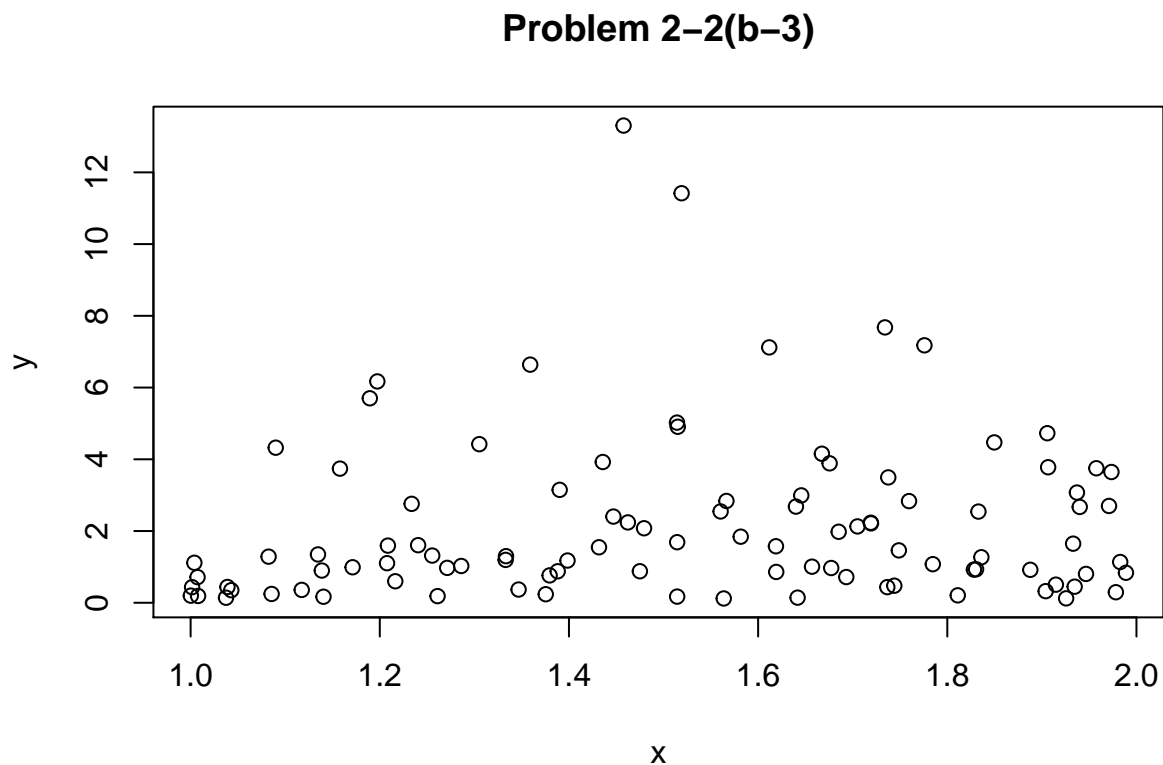
b-1) Read about the random number generator function for exponential distribution, rexp in R help.

b-2) Generate 100 observations of duration data y by i) generating 100 x_i 's from the i.i.d. uniform distribution on the interval (1,2), ii) using $\beta = 0.5$, and iii) and using **rexp** command.

```
set.seed(42)
n <- 100
x <- runif(n, min=1, max=2)
beta <- 0.5
y <- rexp(n, beta)
```

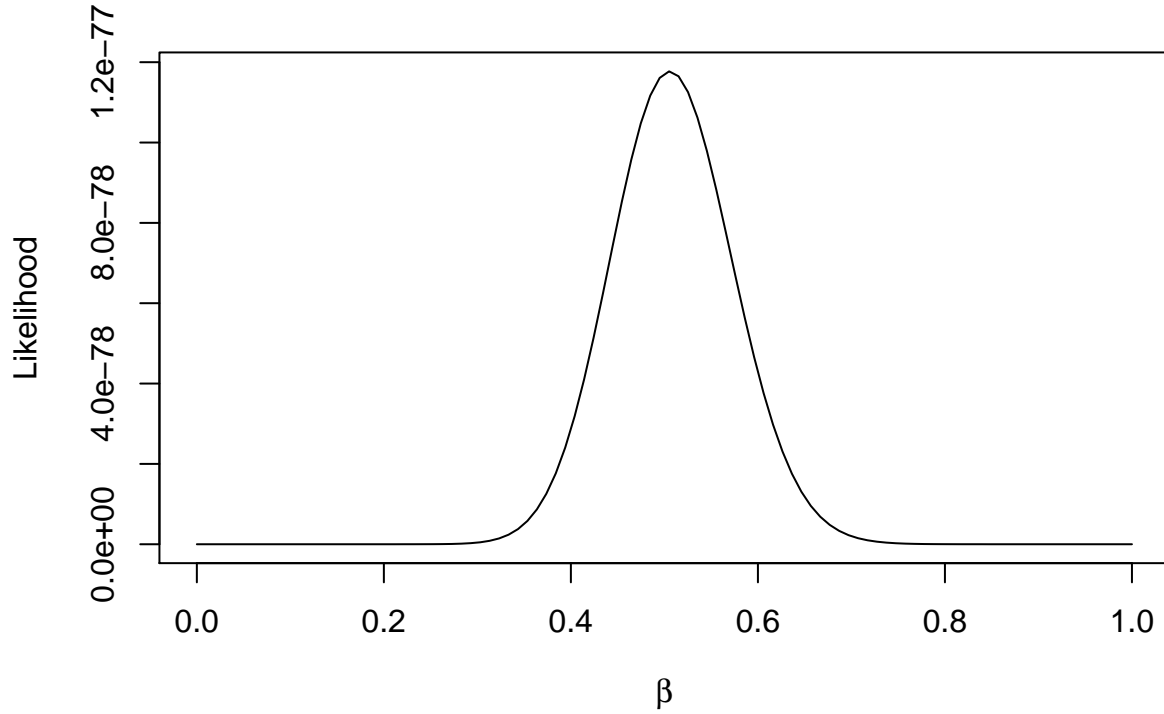
b-3) Plot the data and likelihood.

```
#plot data
plot(x, y, main="Problem 2-2(b-3)")
```



```
# plot likelihood
betaval<- seq(0,1,length=n)
lhood <- c()
for (b in betaval){
  lhood_i <- exp(-b*sum(x))*exp(-sum(y*exp(-b*x)))
  lhood <- c(lhood, lhood_i)
}
invisible(plot(betaval,lhood,type="l", xlab="", ylab="") +
  title(main="Problem 2-2(b-3)", xlab=expression(beta), ylab="Likelihood"))
```

Problem 2-2(b-3)



Problem 3

Consider random variable Y with a Poisson distribution:

$$P(y|\theta) = \frac{\theta^y e^{-\theta}}{y!}, y = 0, 1, 2, \dots, \theta > 0$$

Mean and variance of Y given θ are both equal to θ . Assume that $\sum_{i=1}^n y_i > 1$.

- 1) Write down the likelihood for n independent realizations of Y and then the posterior density of θ under the standard vague prior $p \propto \frac{1}{\theta}$.

Answer

The likelihood can be calculated as follows:

$$L(\theta; y) = \prod_{i=1}^n \frac{\theta^{y_i}}{y_i!} e^{-\theta}$$

We can express $\sum_{i=1}^n y_i$ as $n\bar{y}$, with \bar{y} = mean of y.

$$L(\theta; y) = \frac{\theta^{n\bar{y}} e^{-n\theta}}{\prod_{i=1}^n y_i!}$$

$$L(\theta; y) \propto \theta^{n\bar{y}} e^{-n\theta}$$

The posterior $P(\theta|y)$ can be expressed as likelihood times prior:

$$P(\theta|y) = L(\theta; y) \times P(y)$$

$$P(\theta|y) \propto \theta^{n\bar{y}} e^{-n\theta} \theta^{-1}$$

$$P(\theta|y) \propto \theta^{n\bar{y}-1} e^{-n\theta}$$

2) Construct an asymptotic normal approximation to the posterior by deriving the Hessian of the log posterior and b) the posterior mode.

Answer

The posterior kernel can be rewritten as follows:

$$P(\theta|y) \propto e^{(n\bar{y}-1) \ln \theta - n\theta}$$

The log posterior is then:

$$\ln P(\theta|y) = (n\bar{y} - 1) \ln \theta - n\theta + \text{const.}$$

To calculate the posterior mode, we have to maximize the above function. Let's calculate the first and second derivative:

$$\frac{d \ln P(\theta|y)}{d\theta} = \frac{n\bar{y} - 1}{\theta} - n \qquad \frac{d^2 \ln P(\theta|y)}{d\theta^2} = -\frac{n\bar{y} - 1}{\theta^2}$$

Given that $n\bar{y} > 1$, we know that the second derivative is negative, indicating concavity. We can find the maximum then by setting the first derivative equal to 0 and solving for θ .

$$\frac{n\bar{y} - 1}{\theta} - n = 0$$

The posterior mode is then:

$$\hat{\theta} = \frac{n\bar{y} - 1}{n} = \bar{y} - \frac{1}{n}$$

We can now substitute $\hat{\theta}$ into the second derivative of the log posterior to calculate the Hessian. Take the negative inverse of that result to calculate the variance.

$$H(\hat{\theta}) = \frac{d^2 \ln P(\theta|y)}{d\theta^2}(\hat{\theta}) = -\frac{n\bar{y} - 1}{\frac{(n\bar{y}-1)^2}{n^2}} = -\frac{n^2}{n\bar{y} - 1}$$

$$-(H(\hat{\theta}))^{-1} = \frac{n\bar{y} - 1}{n^2}$$

Thus, we can represent a normal approximation to the posterior as:

$$P(\theta|y) \sim N\left(\frac{n\bar{y} - 1}{n}, \frac{n\bar{y} - 1}{n^2}\right)$$

3) Simulate some data using R (rpois command) for different n=10, 50, 100, 200 and compare graphically the exact gamma posterior density of θ to its normal approximation.

```
lambda <- 3
par(mfrow=c(2,2),bty="l",mar=c(4,2,2,2))
for (n in c(10,50,100,200)){
  # generate true distribution
  set.seed(3)
  y <- rpois(n, lambda)

  # generate posterior, which can be expressed as a gamma
```

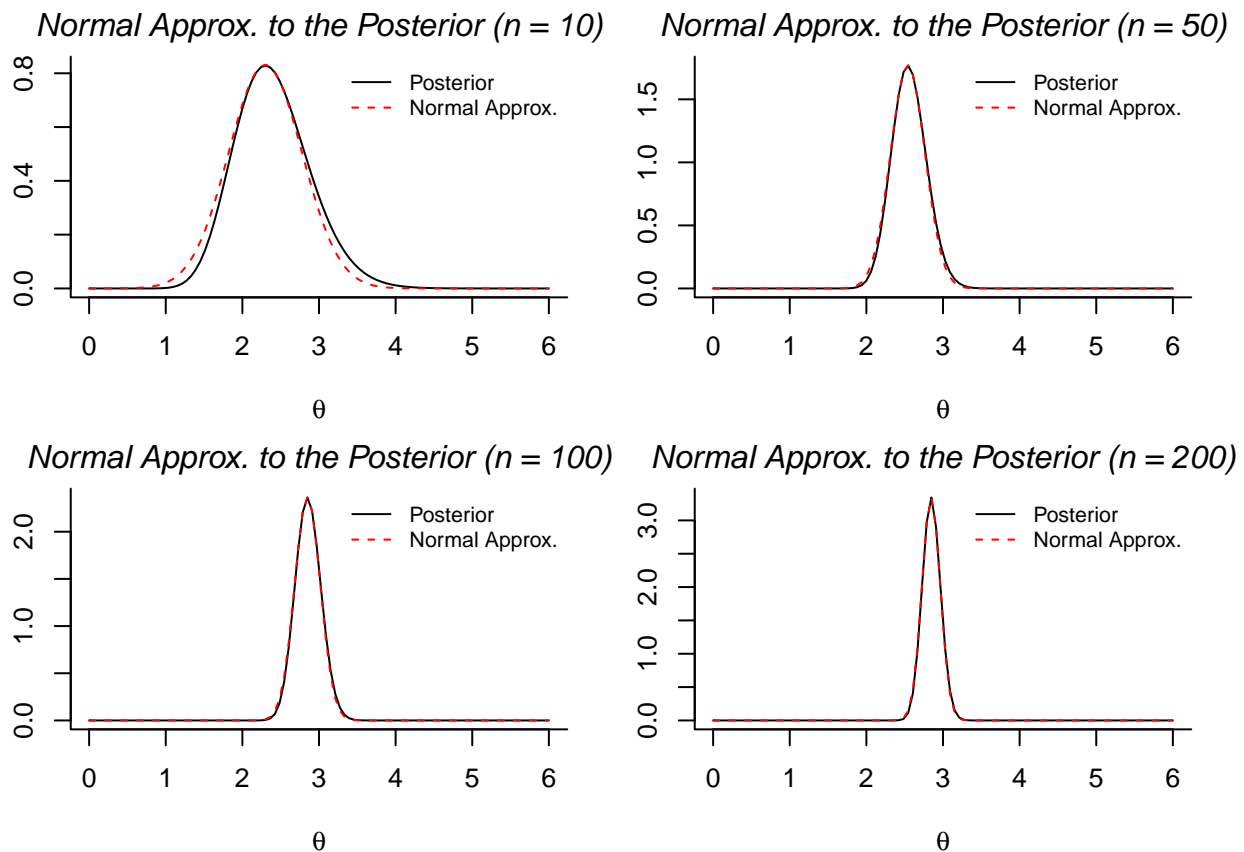
```

thetavals = seq(0,6,length=100)
i <- 1:n
post = dgamma(thetavals,n*mean(y),n)

# generate normal approximation
nmean = (n*mean(y)-1)/n
nsd = sqrt((n*mean(y)-1)/n^2)
norm = dnorm(thetavals, mean=nmean, sd=nsd)

#plot exact posterior and normal
plot(thetavals,post,type='l',ylab="",xlab=expression(theta))+
  lines(thetavals,norm,col='red',lty=2)+
  title(sprintf('Normal Approx. to the Posterior (n = %d)',n), font.main=3)
legend("topright",legend=c('Posterior','Normal Approx.'),cex=0.8,bty="n",lty=c(1,2),col=c("black","red"))
}

```



We see that as n increases, the normal approximation to the posterior gets better and better. Additionally, we see as n increases, the distribution gets closer and closer to the “true” value of the lambda parameter 3.

Problem 4

Consider the following exponential distribution:

$$P(y|\theta) = \theta e^{-\theta y}, \theta > 0, y > 0$$

Assume that y_1, y_2, \dots, y_n are independent random samples from this distribution.

- 1) Show that the gamma distribution $Gamma(\alpha, \beta)$ is a conjugate prior distribution for the exponential distribution.

Answer

Prior can be represented as the following:

$$P(y) = \frac{\beta^\alpha \theta^{\alpha-1} e^{-\beta\theta}}{\Gamma^\alpha}$$

With $\sum_{i=1}^n y_i = n\bar{y}$, with \bar{y} = the mean of y, the likelihood can be expressed as follows:

$$L(\theta; y) = \theta^n e^{-\theta n\bar{y}}$$

Multiplying the likelihood and prior together, we can calculate the posterior:

$$P(\theta|y) = \theta^n e^{-\theta n\bar{y}} \times \frac{\beta^\alpha \theta^{\alpha-1} e^{-\beta\theta}}{\Gamma^\alpha}$$

$$P(\theta|y) = \theta^{n+\alpha-1} e^{-\theta n\bar{y}-\beta\theta} \times \frac{\beta^\alpha}{\Gamma^\alpha}$$

$$P(\theta|y) \propto \theta^{n+\alpha-1} e^{-\theta n\bar{y}-\beta\theta}$$

This posterior can be represented in the family of gamma distributions as follows, thus proving that the gamma distribution is a conjugate prior for the exponential:

$$Gamma(n + \alpha, n\bar{y} + \beta)$$

- 2) Derive the maximum likelihood estimator for θ .

Answer

To calculate the MLE, we need to find the maximum of the log-likelihood. Let's take the log of the likelihood, which was calculated above:

$$\ln L(\theta; y) = -\theta n\bar{y} + n \ln \theta$$

Take the first derivative and second derivative:

$$\frac{d \ln L(\theta; y)}{d\theta} = -n\bar{y} + \frac{n}{\theta} \qquad \frac{d^2 \ln L(\theta; y)}{d\theta^2} = -\frac{n}{\theta^2}$$

The second derivative is negative, indicating concavity. This means we can set the first derivative equal to 0 and obtain the maximum. Doing this returns the maximum likelihood estimator:

$$\hat{\theta} = \frac{n}{n\bar{y}} = \frac{1}{\bar{y}}$$