XIAO_Lin_Solutions_lab8

Lin Xiao

November 2, 2015

Today we will be working with another very popular conjugate model: the Gamma-Exponential. In this case, we are interested in determining whether the true parameter rate of decline of an engine of a Toyota Hilux is $\theta = 0.5$. For this purpose, conduct the following investigation:

1. Start by writing out three expressions: 1) An Exponential likelihood with data Y and rate θ 2) A Gamma prior with parameters (a,b) 3) The posterior for θ. Were you able to confirm conjugacy?

   The exponential likelihood with data Y and rate θ

   $$f(y|\theta) = \begin{cases} \theta e^{-\theta y} & y > 0 \\ 0 & y \leq 0 \end{cases}$$

   Gamma prior with parameter (a,b)

   $$p(\theta) = \frac{b^a}{\Gamma(a)} \theta^{a-1} e^{-b\theta} \sim Gamma(a, b)$$

   For all $y_i > 0$ we have the posterior for θ

   $$\begin{aligned} p(\theta|y) &= \frac{\prod_{i=1}^{n} f(y_i|\theta)p(\theta)}{f(y)} \\ &= \frac{\prod_{i=1}^{n} f(y_i|\theta)p(\theta)}{\int_0^\infty \prod_{i=1}^{n} f(y_i|\theta)p(\theta)d\theta} \\ &= \frac{(b + \sum y_i)^{a+n}}{\Gamma(a+n)} \theta^{(a+n)-1} e^{-(b+\sum y_i)\theta} \\ &\sim Gamma(a+n, b+\sum y_i) \end{aligned}$$

   We can see from the forms of posterior and prior that they are conjugate distributions.

2. Simulate some data using the rexp function of size n = 100 and the rate equals to 0.5. Remember to set.seed(224) so that I can replicate your results.

```r
library(BB)
set.seed(224)
dat <- rexp(100, 0.5)
```

3. Since the Toyota Hilux is practically as old as time, there are many, many simulation studies that have already been done to assess its engine. Assume that we believe that the median value of $\theta = 0.1$ and 0.5 is closer to the θ_{80th}. Using your knowledge of BBsolve, derive the a and b that satisfy these conditions. Intuitively explain what you have done here in terms of finding a and b

```r
fn = function(x){qgamma(c(0.5,0.8),x[1],x[2])-c(0.1,0.5)}
a <- BBsolve(c(1,1),fn)$par[1];a
```

```
##   Successful convergence.
```

```
## [1] 0.3596391
```

```r
b <- BBsolve(c(1,1),fn)$par[2];b
```

```
##   Successful convergence.
```

```
## [1] 1.144024
```

What I have done with BBsolve is solving the following calculus problems:

$$\int_0^{0.1} \frac{b^a}{\Gamma(a)}\theta^{a-1}e^{-b\theta} = 0.5$$

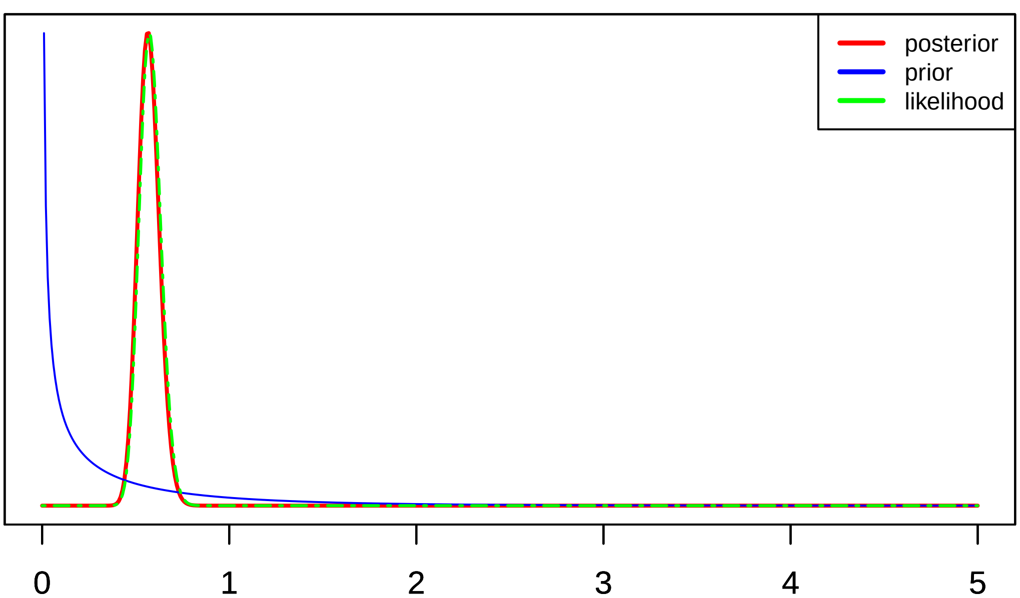
$$\int_0^{0.5} \frac{b^a}{\Gamma(a)}\theta^{a-1}e^{-b\theta} = 0.8$$

4. Using the values of a and b create a plot of your prior, likelihood and posterior in the same window. Recall that you must turn the Y axis off for this plot to be reasonable. Describe the features that you see. Note that you will need to write your own likelihood function in this case. Be sure that your plot highlights the important aspects of the data.

```r
# Generate a sequence of theta
theta <- seq(0, 5, by = 0.01)
n <- 100

# Calculate the likelihood values
lhd <- theta^n*exp(-sum(dat)*theta)

# Plotting order: posterior, prior and likelihood
par(yaxt="n",  ann=FALSE)
plot(theta, dgamma(theta,a+n,sum(dat)+b),
     type="l", col="red",lty=1, lwd=2)
par(new = T)
plot(theta, dgamma(theta,a,b),type="l",col="blue")
par(new = T)
plot(theta, lhd, col="green",type="l",lty=6, lwd=1.5)
legend("topright", c("posterior","prior","likelihood"),lty=c(1,1,1),
       lwd=c(2.5,2.5,2.5), col=c("red", "blue", "green"),cex=0.75)
```



From the plot we can see that the shapes of posterior and likelihood are almost the same, which means that the prior is considered weakly-informative.

5. Generate a sample of size T = 1000 from your posterior using the appropriate function and based on this provide a 95% credible interval (remember to use xtable). What is your conclusion to the original question of interest?

```
library(xtable)
Cred_I <- quantile(rgamma(1000,a+n,sum(dat)+b), c(0.025,0.975))


dat_CI <- data.frame(Cred_I)
names(dat_CI) <- c("Credible Interval")
tb <- xtable(dat_CI)
digits(tb)[2] <- 6
print(tb, comment=F)
```

	Credible Interval
2.5%	0.466840
97.5%	0.686668

We can conclude that, since 0.5 is inside the credible interval, then we do not have significance evidence to reject that the true parameter rate of decline of an engine of a Toyota Hilux is $theta = 0.5$.