Homework 1 Solution
of
STAT 632 Bayesian Statistics

Yanxun Xu

September 15, 2015

1 Problem 1

Solution:

The data likelihood of $X = [11.0, 11.5, 11.7, 11.1, 11.4, 10.9]$ is

$$p(X|\theta) = 1, \theta - \frac{1}{2} \leq x_i \leq \theta + \frac{1}{2},$$

where $\max\{X\} + \frac{1}{2} \leq \theta \leq \min\{X\} + \frac{1}{2}$, that is, $11.2 \leq \theta \leq 11.4$.

The prior distribution of θ is

$$p(\theta) = 1, 10 \leq \theta \leq 20.$$

The posterior distribution of θ is

$$p(\theta|X) \propto p(X|\theta)p(\theta) \propto 1.$$

Because of the available data, we know that $11.2 \leq \theta \leq 11.4$, so the normalizing constant equals to $\frac{1}{11.4-11.2} = 5$. Therefore,

$$p(\theta|X) = 5.$$

2 Problem 2

Solution:

(a) According to the condition $p(x, y, z) \propto f(x, z)g(y, z)h(z)$ as well as Bayes' rule, we have

$$\begin{aligned} p(x|y,z) &= \frac{p(x,y,z)}{g(y,z)} \\ &\propto f(x,z)h(z) \\ &\propto f(x,z), \end{aligned}$$

where $p(x|y, z)$ is a function of x and z.

(b) According to the condition $p(x, y, z) \propto f(x, z)g(y, z)h(z)$ as well as Bayes' rule, we have

$$\begin{aligned} p(y|x,z) &= \frac{p(x,y,z)}{f(x,z)} \\ &\propto g(y,z)h(z) \\ &\propto g(y,z), \end{aligned}$$

where $p(y|x,z)$ is a function of y and z.

(c) To show that X and Y are conditionally independent given Z, we should prove that $p(x,y|z) = p(x|z)p(y|z)$.

$$\begin{aligned}
p(x,y|z) &= \frac{p(x,y,z)}{h(z)} \\
&= \frac{p(x|y,z)g(y,z)}{h(z)} \\
&= p(x|y,z)p(y|z) \\
&\propto f(x,z)p(y|z) \text{ \textit{(according to the result of (a))}} \\
&\propto p(x|z)p(z)p(y|z) \\
&\propto p(x|z)p(y|z)
\end{aligned}$$

3 Problem 3

Solution:

The data likelihood of X is

$$p(X|\theta) = \prod_{i=1}^{n} \theta \exp(-\theta x_i) = \theta^n \exp(-\theta n\bar{x}),$$

where $n = 20$ and $\bar{x} = 3.8$ in this case.

The prior distribution of θ is

$$p(\theta) = Gamma(\theta; \alpha, \beta),$$

where $\alpha = 0.04$ and $\beta = 0.2$ in this case.

The posterior distribution of θ is

$$\begin{aligned}
p(\theta|X) &\propto p(X|\theta)p(\theta) \\
&= \theta^n \exp(-\theta n\bar{x}) Gamma(\theta; \alpha, \beta) \\
&= \theta^n \exp(-\theta n\bar{x}) \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} \exp(-\theta\beta) \\
&\propto \theta^{\alpha+n-1} \exp(-\theta(\beta + n\bar{x})) \\
&= Gamma(\theta; \alpha + n, \beta + n\bar{x}) \\
&= Gamma(\theta; 20.04, 76.2)
\end{aligned}$$

4 Problem 4

Solution:

The data likelihood of r follows a negative binomial distribution with an unknown parameter p, given by

$$p(r|p) = (\begin{array}{c} k+r-1 \\ k \end{array})(1-p)^r p^k.$$

We assume that the prior distribution of p follows A beta distribution with parameters α and β, given by

$$p(p) = beta(\alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}.$$

The posterior distribution of p is

$$\begin{aligned} p(p|r) &\propto p(r|p)p(p) \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}\left(\begin{array}{c} k+r-1 \\ k \end{array} \right)(1-p)^r p^k \\ &= c(\alpha,\beta,k,r)p^{\alpha+k-1}(1-p)^{\beta+r-1} \\ &= beta(\alpha+k,\beta+r). \end{aligned}$$

The third to last line says that $p(p|r)$ is proportional to $p^{\alpha+k-1}(1-p)^{\beta+r-1}$, which means that it has the same shape as $beta(\alpha+k,\beta+r)$. But we also know that $p(p|r)$ and the beta density must both integrate to 1, and therefore they are share the same scale. Thus, the family of beta distributions is a conjugate family of prior distributions for samples from a negative binomial distribution with an unknown parameter p.

5 Problem 5

Solution:

(a) The joint distribution of $P(Y_1 = y_1, \ldots, Y_{100} = y_{100}|\theta)$ can be written as,

$$P(Y_1 = y_1, \ldots, Y_{100} = y_{100}|\theta) = \theta^{\sum_{i=1}^{100} y_i}(1-\theta)^{100-\sum_{i=1}^{100} y_i},$$

which can also be written as,

$$P(\sum_{i=0}^{100} Y_i = y|\theta) = \left(\begin{array}{c} 100 \\ y \end{array} \right)\theta^y(1-\theta)^{100-y}.$$

(b) The data likelihood of $\sum_{i=1}^{100} Y_i = 57$ for each of those 11 values of θ and the corresponding plot are given in Table 1 and Figure 1 (Upper), respectively.

Table 1: The data likelihood for each parameter

θ	0.0	0.1	0.2	0.3	0.4	0.5
$P(\sum_{i=0}^{100} Y_i = 57\|\theta)$	0.000	4.107×10^{-31}	3.738×10^{-16}	1.307×10^{-8}	2.286×10^{-4}	3.007×10^{-2}
θ	0.6	0.7	0.8	0.9	1.0	
$P(\sum_{i=0}^{100} Y_i = 57\|\theta)$	6.673×10^{-2}	1.853×10^{-3}	1.004×10^{-7}	9.396×10^{-18}	0.000	

(c) The posterior distribution for each of those 11 values of θ and the corresponding plot are given in Table 2 and Figure 1 (Lower), respectively.

Table 2: The posterior distribution for each parameter

θ	0.0	0.1	0.2	0.3	0.4	0.5
$P(\sum_{i=0}^{100} Y_i = 57\|\theta)$	0.000	4.148×10^{-29}	3.776×10^{-14}	1.320×10^{-6}	2.309×10^{-2}	3.037×10^{0}
θ	0.6	0.7	0.8	0.9	1.0	
$P(\sum_{i=0}^{100} Y_i = 57\|\theta)$	6.740×10^{0}	1.872×10^{-1}	1.014×10^{-5}	9.490×10^{-16}	0.000	

(d) The posterior density $p(\theta)P(\sum_{i=0}^{100} Y_i = y|\theta)$ as a function of θ is shown in Figure 2 (Upper).

(e) The posterior density $P(\theta|\sum_{i=0}^{100} Y_i = y)$ as a function of θ is shown in Figure 2 (Lower). Figure 1 and Figure 2 tells us that a discrete or continuous uniform prior distribution gives a posterior distribution that is proportional to the sampling probability, i.e., data likelihood.

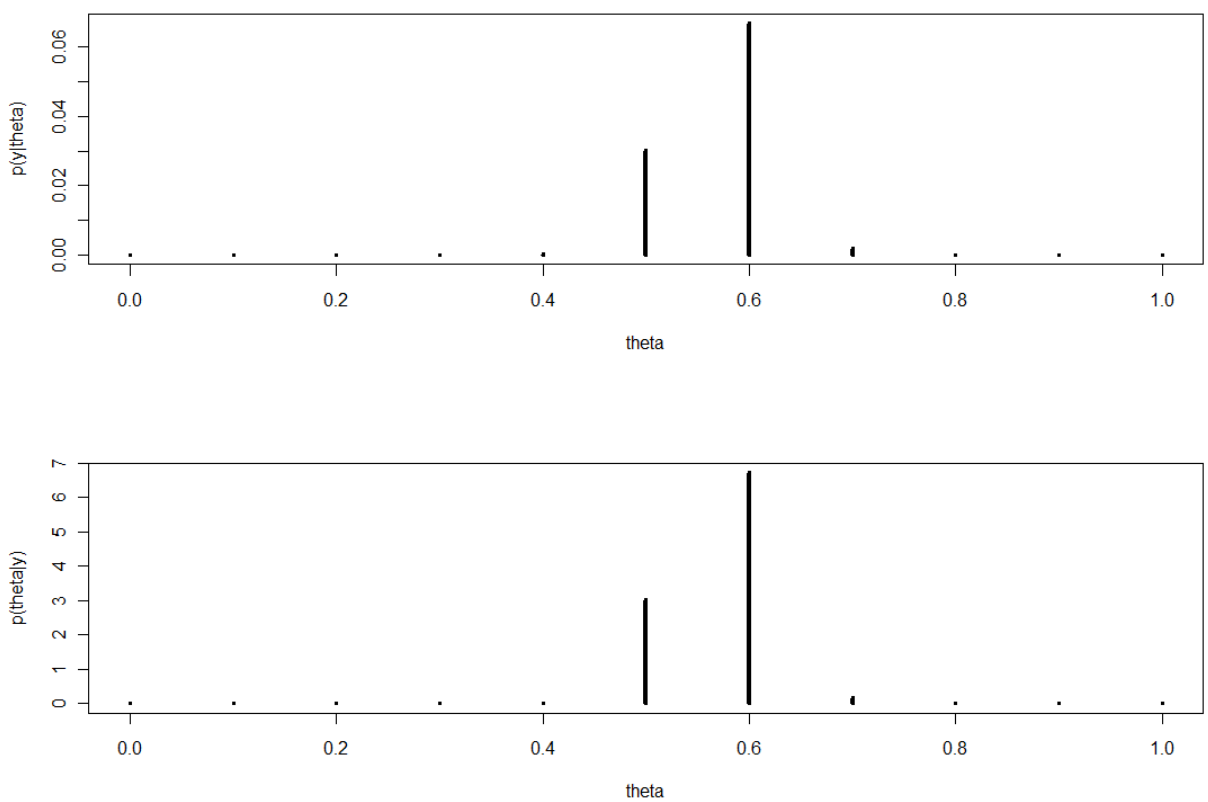

Figure 1: The data likelihood (Upper) and the posterior distribution (Lower) as a function of θ

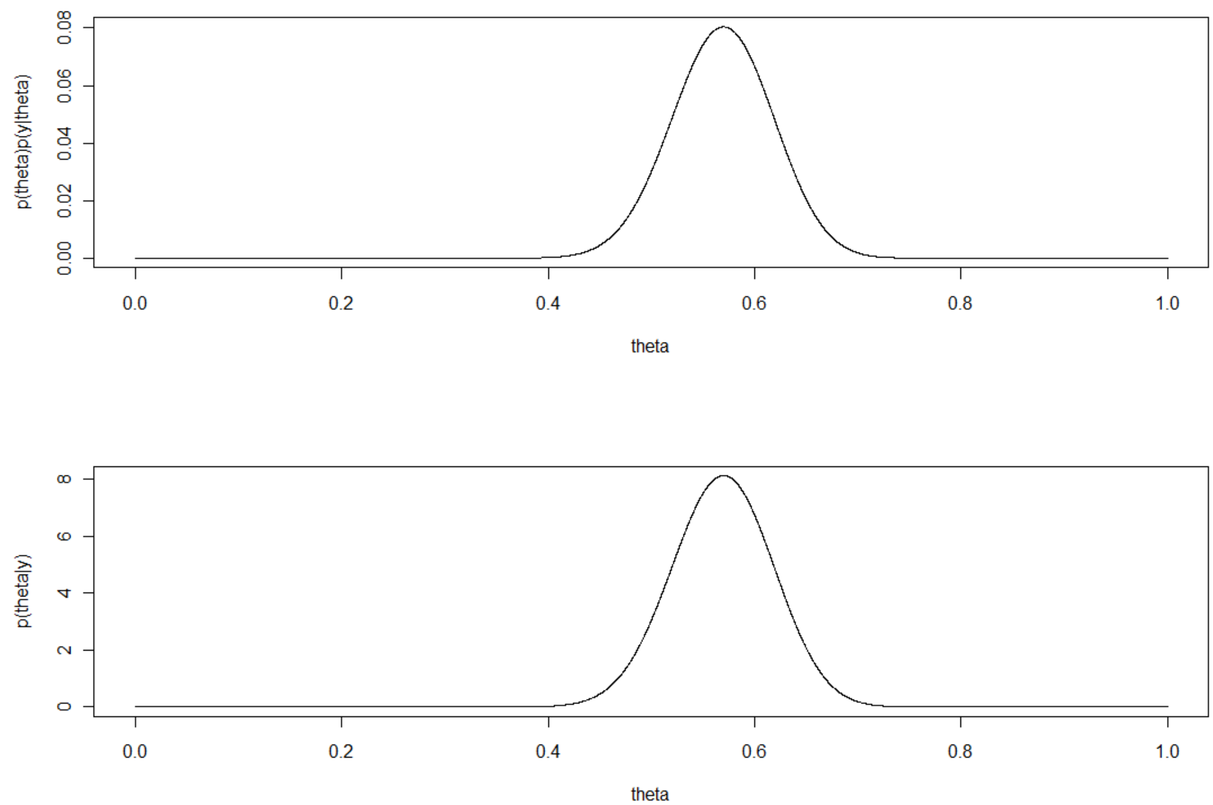

Figure 2: The posterior density $p(\theta)P(\sum_{i=0}^{100} Y_i = y|\theta)$ (Upper) and the posterior distribution density $P(\theta|\sum_{i=0}^{100} Y_i = y)$ (Lower) as a function of θ

6 Problem 6

Solution:

Using the following code, we are able to find the corresponding a, b values and compute $P(\theta > 0.5)|\sum_{i=0}^{100} Y_i = 57$, displaying with a contour plot, as shown in Figure 3. The plot indicates that low values of n_0 (weak prior beliefs) or high prior expectations θ_0 are generally 90% or more certain that the approval ratings θ is more than 0.5.

```
library(ggplot2)
theta_0 = c(1:9) / 10
n_0 = c(1, 2, 8, 16, 32)
for(i in 1:length(theta_0)){
  for(j in 1:length(n_0)){
    a[(i - 1) * length(n_0) + j] = theta_0[i] * n_0[j]
    b[(i - 1) * length(n_0) + j] = (1 - theta_0[i]) * n_0[j]
    prob[(i - 1) * length(n_0) + j] =
      sum(rbeta(1000, theta_0[i] * n_0[j] + y, (1 - theta_0[i]) * n_0[j] + n - y) > 0.5) / 1000
  }
}
qplot(a + b, a / (a + b), prob) + geom_contour(aes(colour = ..level..))
```

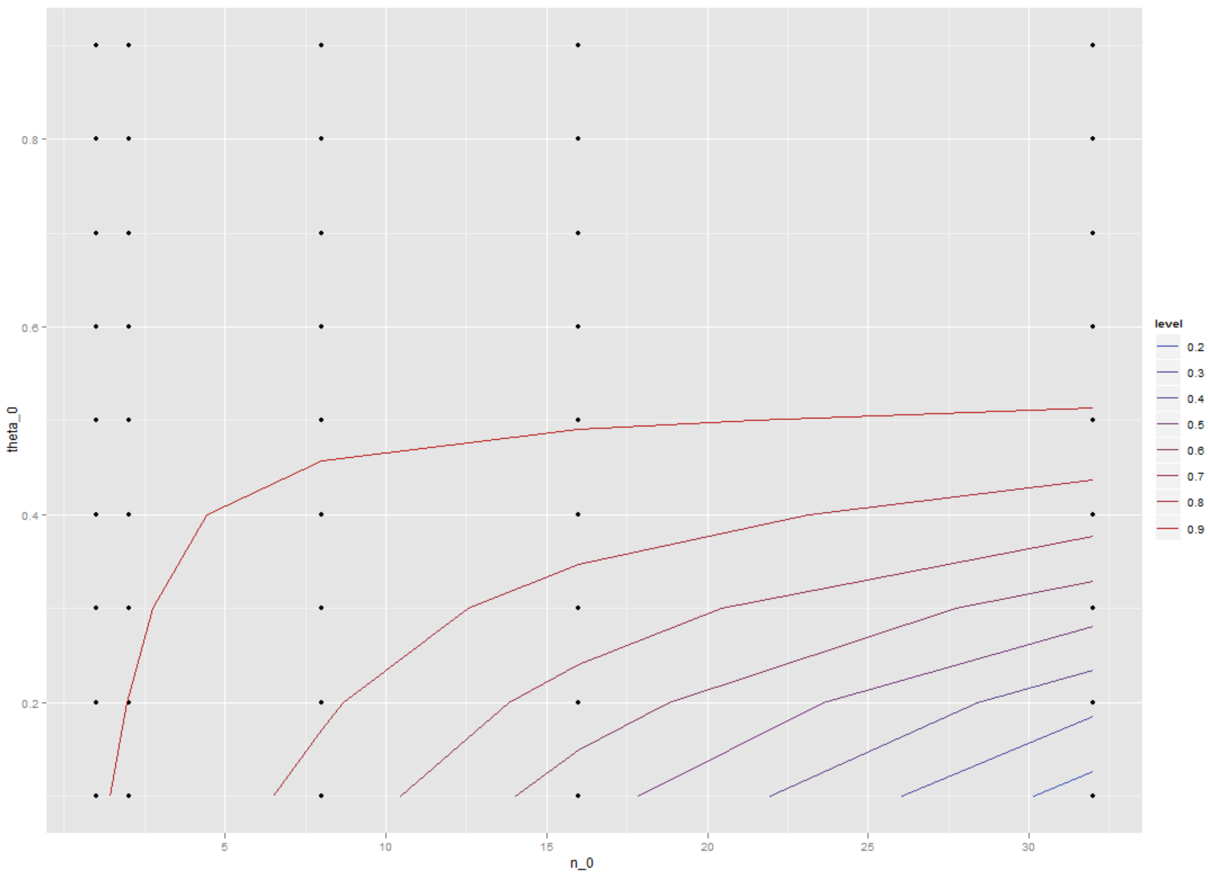

Figure 3: A contour plot of posterior quantities under different beta prior distributions