Bayesian Models 1

Charlie Qu

June 14, 2017

Outline

This is a first short writing about Bayesian statistics. I am going to work through an example of performing posterior inference for data that have a binomial distribution.

1. Load in the data;
2. Perform some exploratory analysis;
3. Perform posterior inference using a discrete prior;
4. Perform posterior inference using a beta prior.

Suppose we want to estimate ??, the proportion of animals that can fly, from a random sample of animals.Before we start looking at the data, let's think about a prior for ??.

(1) What values can it take on?
(2) What's a reasonable prior mean?
(3) How uncertain are we about our prior mean?

Suppose we want to estimate ??, the proportion of animals that can fly, from a random sample of animals.

We'll use the animals data set, which measures characteristics of 20 species.

It is available in the **cluster** package, one of the basic packages that comes with R.

```
library(cluster)
```

```
## Warning: package 'cluster' was built under R version 3.3.3
```

```
data(animals)
```

Description of the Data

The animals data includes indicators:

(1) war, warm-blooded;
(2) fly, able to fly;
(3) ver, vertebrate;
(4) end, endangered;
(5) gro, live in groups;
(6) hai, have hair.

```
help(animals)
```

```
## starting httpd help server ... done
```

Preparing the Data

The indicators are coded to take on:2, if the animal has the characteristic;1, otherwise.

We want to replace the 2's and 1's with 1's and 0's.

```r
animals <- apply(animals, c(1, 2),
                 function(x) {x - 1})
```

Describing the Data

Since we're interested in ??, the proportion of animals that can fly, we should look at the number of animals in the sample that can fly.

```r
sum(animals[, "fly"])
```

```
## [1] 4
```

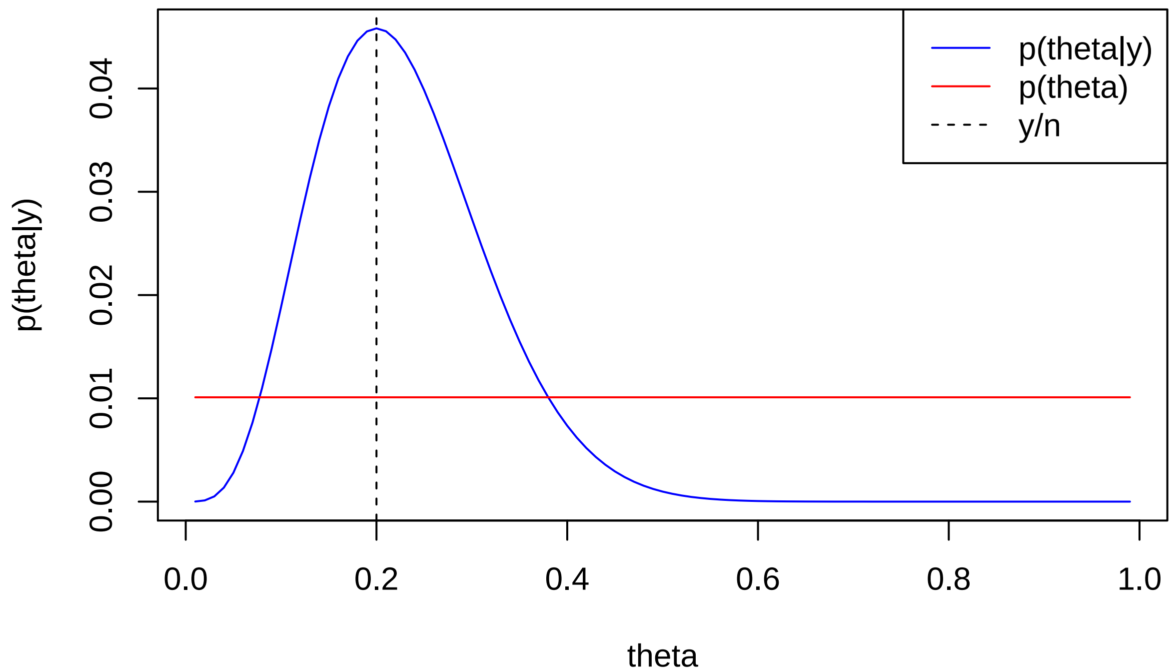
So we have y=4 from a Binomial(n=20, ??) distribution.

Inference Under a Discrete Prior

Suppose we believed:

(1) ??={0.01,.,0.99};
(2) p(??)=1/99 for each ???????. This prior encodes the belief that it is impossible for: (a)All animals to fly, i.e. ?????1; (b)No animals to fly, i.e. ?????0. This is a uniform prior on ??. We can compute the posterior distribution using Bayes' rule.

```r
y <- sum(animals[, "fly"])
n <- nrow(animals)

Theta.1 <- seq(0.01, 0.99, by = 0.01)
py.theta.1 <- dbinom(y, n, Theta.1)
ptheta.1 <- rep(1/length(Theta.1), length(Theta.1))
pytheta.1 <- py.theta.1*ptheta.1

ptheta.y.1 <- pytheta.1/(sum(pytheta.1))
plot(Theta.1, ptheta.y.1, type = "l",
     col = "blue",
     xlab = "theta",
     ylab = "p(theta|y)",
     cex.lab = 1, cex.axis = 1)

lines(Theta.1, ptheta.1, col = "red", lty = 1)

abline(v = mean(animals[, "fly"]), lty = 2)

legend("topright", lty = c(1, 1, 2),
       col = c("blue", "red", "black"),
       legend = c("p(theta|y)", "p(theta)",
                  "y/n"))
```



We can compute the posterior mean easily.

```
etheta.y.1 <- sum(ptheta.y.1*Theta.1)
etheta.y.1
```

```
## [1] 0.2272727
```

We can also compute the posterior variance easily.

```
vtheta.y.1 <- sum(ptheta.y.1*(Theta.1 - etheta.y.1)^2)
vtheta.y.1
```

```
## [1] 0.007635645
```

Inference Under a Continuous Prior

Suppose we believed: (1) ??=(0,1); (2) p(??)=1 for each ???????. This is also uniform prior.

This prior corresponds to the Beta(1, 1) distribution.

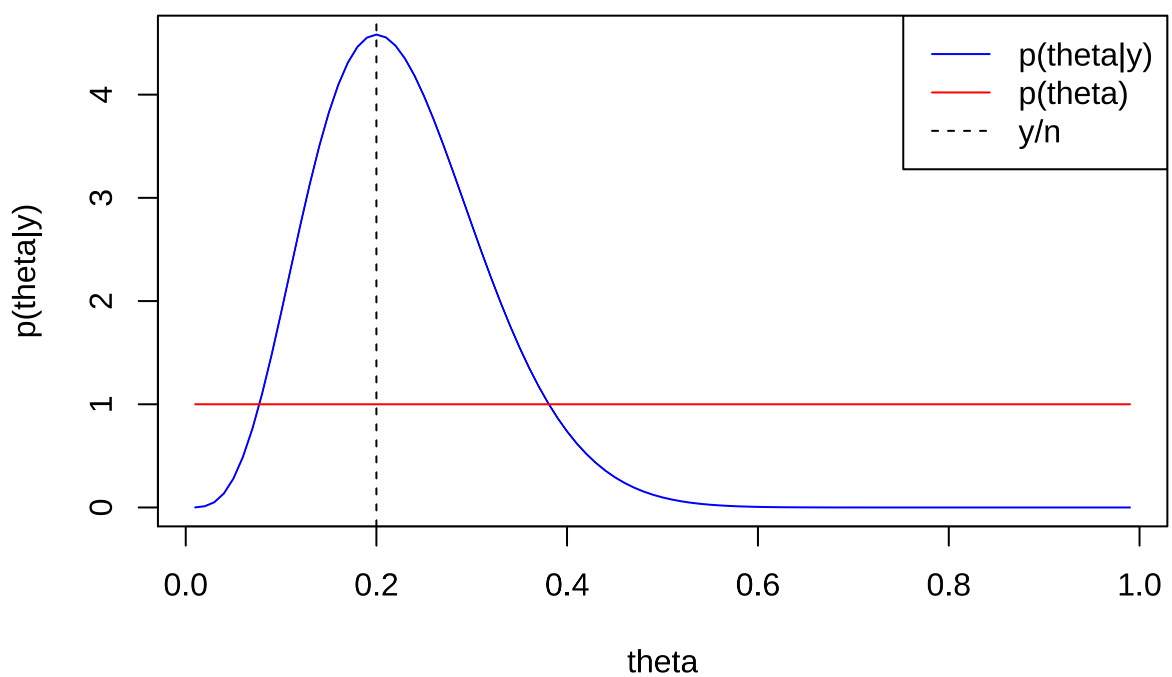
As you'll learn later, the beta distribution is conjugate for the binomial sampling model.

This means that we know that the posterior distribution of ?? will be a beta distribution, with parameters we have closed form solutions for!

??|y???B(1+y,1+n???y)

```
atheta.y.2 <- 1 + y
btheta.y.2 <- 1 + n - y
Theta.2 <- seq(0.01, 0.99, by = 0.01)
```

```r
ptheta.2 <- dbeta(Theta.2, 1, 1)
ptheta.y.2 <- dbeta(Theta.2, atheta.y.2,
                    btheta.y.2)

plot(Theta.2, ptheta.y.2, type = "l",
     col = "blue",
     xlab = "theta",
     ylab = "p(theta|y)",
     cex.lab = 1, cex.axis = 1)

lines(Theta.2, ptheta.2, col = "red", lty = 1)
abline(v = mean(animals[, "fly"]), lty = 2)
legend("topright", lty = c(1, 1, 2),
       col = c("blue", "red", "black"),
       legend = c("p(theta|y)", "p(theta)",
                  "y/n"))
```



We can directly compute the posterior mean from what we know about the beta distribution.

```r
etheta.y.2 <- atheta.y.2/(atheta.y.2 + btheta.y.2)
etheta.y.2
```

```
## [1] 0.2272727
```

We can also compute the posterior variance.

```r
vtheta.y.2 <- atheta.y.2*btheta.y.2/((atheta.y.2 + btheta.y.2)^2*(atheta.y.2 + btheta.y.2 + 1))
vtheta.y.2
```

```
## [1] 0.007635645
```

Note that we can compute the mode and any other feature of the Beta distribution that can be written as a function of its parameters.

Inference Under Another Continuous Prior

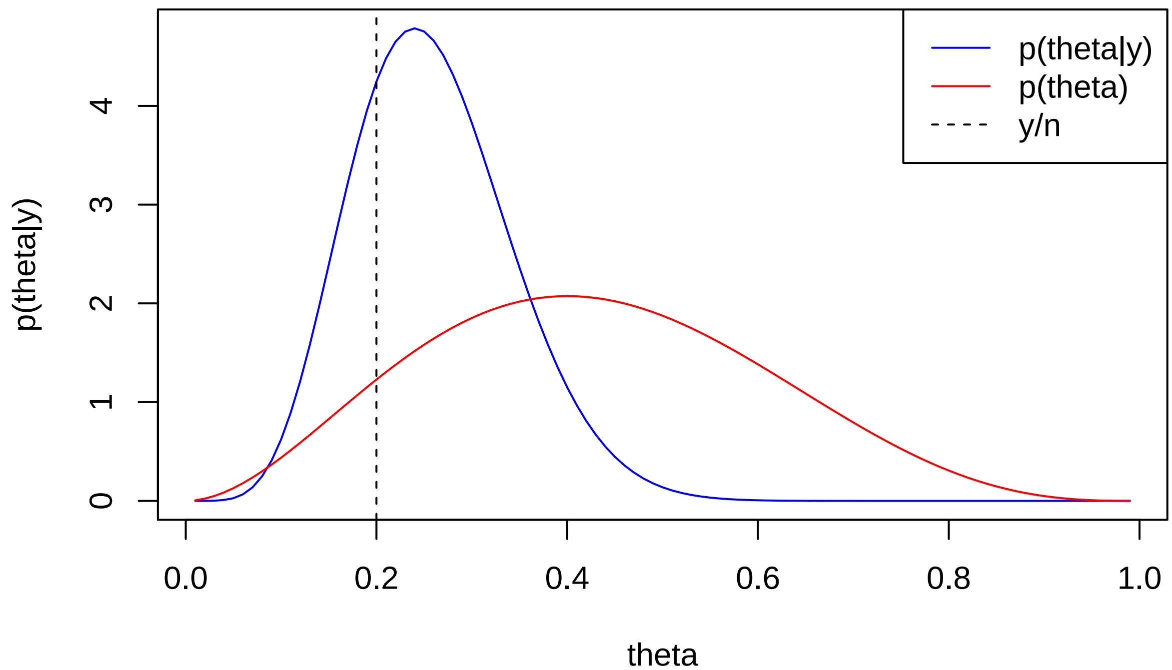
Suppose we believed:

(1) ??=(0,1);
(2) E[??]=0.43, i.e. about 2/5 of animals fly;
(3) V[??]=0.03, i.e. we're pretty certain ??'s around 0.43.

This corresponds to a beta distribution with a=3 and b=4.

Again, we can use the fact that the beta distribution is conjugate for the binomial sampling model to compute the parameters of the posterior distribution.

```
atheta.y.3 <- 3 + y
btheta.y.3 <- 4 + n - y
Theta.3 <- seq(0.01, 0.99, by = 0.01)
ptheta.3 <- dbeta(Theta.3, 3, 4)
ptheta.y.3 <- dbeta(Theta.3, atheta.y.3,
                    btheta.y.3)

plot(Theta.3, ptheta.y.3, type = "l",
     col = "blue",
     xlab = "theta",
     ylab = "p(theta|y)",
     cex.lab = 1, cex.axis = 1)

lines(Theta.3, ptheta.3, col = "red", lty = 1)
abline(v = mean(animals[, "fly"]), lty = 2)
legend("topright", lty = c(1, 1, 2),
       col = c("blue", "red", "black"),
       legend = c("p(theta|y)", "p(theta)",
                  "y/n"))
```



Again, we can directly compute the posterior mean from what we know about the beta distribution.

```
etheta.y.3 <- atheta.y.3/(atheta.y.3 + btheta.y.3)
etheta.y.3
```

```
## [1] 0.2592593
```

We can also compute the posterior variance.

```
vtheta.y.3 <- atheta.y.3*btheta.y.3/((atheta.y.3 + btheta.y.3)^2*(atheta.y.3 + btheta.y.3 + 1))
vtheta.y.3
```

```
## [1] 0.006858711
```

Again, we can compute the mode and any other feature of the Beta distribution that can be written as a function of its parameters.