Fall 2019 Jingchen (Monika) Hu

MATH 347 Homework 2 (Total 30 points)

Due: Tuesday 9/24, at the beginning of the class

Name: Shashank Sule

- **Print out this cover page and staple with your homework.**
- **Show all work. Incomplete solutions will be given no credit.**
- **You may prepare either hand-written or typed solutions, but make sure that they are legible. Answers that cannot be read will be given no credit.**
- **R graphical outputs must be printed instead of hand-drawn.**

1. (*6 points; 2 points each part*)
Recall the Tokyo Express dining preference example covered in class. Suppose the Tokyo Express owner in the college town gives another survey to a different group of students. This time, he gives to 30 students and receive 10 of them saying Friday is their preferred day to eat out. Use the owner's prior (restated below) and calculate the following posterior probabilities.

$$\begin{aligned} p &= \{0.3, 0.4, 0.5, 0.6, 0.7, 0.8\} \\ \pi_{owner}(p) &= (0.125, 0.125, 0.250, 0.250, 0.125, 0.125) \end{aligned}$$

   (a) The probability that 30% of the students prefer eating out on Friday.

   (b) The probability that more than half of the students prefer eating out on Friday.

   (c) The probability that between 20% and 40% of the students prefer eating out on Friday.

2. (*8 points; 2 points each part*)
Revisit the figure in lecture slides page 23, where nine different Beta curves are displayed. In the context of Tokyo Express customers' dining preference example where p is the proportion of students preferring Friday, interpret the following prior choices in terms of the opinion of p. For example, Beta$(0.5, 0.5)$ represents the prior belief the extreme values $p = 0$ and $p = 1$ are more probable and $p = 0.5$ is the least probable. In the customers' dining preference example, specifying a Beta$(0.5, 0.5)$ prior indicates the owner thinks the students' preference of dining out on Friday is either very strong or very weak.

   (a) Interpret the Beta$(1, 1)$ curve.

   (b) Interpret the Beta$(0.5, 1)$ curve.

   (c) Interpret the Beta$(4, 2)$ curve.

   (d) Compare the opinion about p expressed by the two Beta curves: Beta$(4, 1)$ and Beta$(4, 2)$.

3. (*8 points; 2 points each part*)
Use any of the functions from this list i) `dbeta()`, ii) `pbeta()`, iii) `qbeta()`, iv) `rbeta()`,

v) `beta_area()`, and vi) `beta_quantile()` to answer the following questions about Beta probabilities.

(a) The density of Beta$(0.5, 0.5)$ at $p = \{0.1, 0.5, 0.9, 1.5\}$.

(b) The probability $P(0.2 \leq p \leq 0.6)$ if $p \sim$ Beta$(6, 3)$.

(c) The quantile of the Beta$(10, 10)$ distribution at the probability values in the set $\{0.1, 0.5, 0.9, 1.5\}$.

(d) A sample of 100 random values from Beta$(4, 2)$.

4. (*4 points*)
If the proportion has a Beta(a, b) prior and one observes Y from a Binomial distribution with parameters n and p, then if one observes $Y = y$, then the posterior density of p is Beta$(a + y, b + n - y)$.

Recall that the mean of a Beta(a, b) random variable following is $\frac{a}{a+b}$. Show that the posterior mean of $p \mid Y = y \sim$ Beta$(a + y, b + n - y)$ is a weighted average of the prior mean of $p \sim$ Beta(a, b) and the sample mean $\hat{p} = \frac{y}{n}$. Find the two weights and explain their implication for the posterior being a combination of prior and data. Comment how Bayesian inference allows collected data to sharpen one's belief from prior to posterior.

5. (*4 points*)
Derivation exercise of the Beta posterior. If the proportion has a Beta(a, b) prior and one samples Y from a Binomial distribution with parameters n and p, then if one observes $Y = y$, then the posterior density of p is Beta$(a + y, b + n - y)$. You need to perform the complete derivation, i.e. keep the constants.

Solutions to Homework 1

Shashank Sule

9/10/2019

Problem 1

Calculating the posterior distribution:

```
##     p Prior    Posterior
## 1 0.3 0.125 4.469407e-01
## 2 0.4 0.125 3.636649e-01
## 3 0.5 0.250 1.766878e-01
## 4 0.6 0.250 1.261301e-02
## 5 0.7 0.125 9.342907e-05
## 6 0.8 0.125 1.068012e-07
```

a. Calculating the probability that 30% of the students prefer eating out on Friday:

```
## Loading required package: dplyr

## Warning: package 'dplyr' was built under R version 3.4.4

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

## Warning: package 'bindrcpp' was built under R version 3.4.4

## [1] 0.4469407
```

b. The probability that more than half the students prefer eating out on a Friday:

```
filtered <- vals %>% filter(p>0.5)
sum(filtered$Posterior)
```

```
## [1] 0.01270655
```

c. The probability that between 20% and 40% of the students prefer eating out on a Friday:

```r
filtered <- vals %>% filter(p>=0.2 & p<=0.4)
sum(filtered$Posterior)
```

```
## [1] 0.8106056
```

Problem 2

a. The owner thinks that every proportion is equally likely, i.e he can't really tell what the customers' preferences are.
b. The owner thinks that the customers don't like eating out on Fridays.
c. The owner thinks that about three-fourths of his customers like eating out on Fridays.
d. According to Beta(4,1), the owner thinks that his customers highly prefer eating out on Friday and that a higher proportion is always more likely than a lower proportion. On the other hand, according to Beta(4,2), the owner thinks that about three-fourths of his customers like eating out on Fridays, and that between any two proportions, the likelier one is the one that is closest to 0.75.

Problem 3

a) Density of Beta(0.5,0.5) at $p = \{0.1, 0.5, 0.9, 1.5\}$

```
##     p      dens
## 1 0.1 1.0610330
## 2 0.5 0.6366198
## 3 0.9 1.0610330
## 4 1.5 0.0000000
```

b) The probability $P(0.2 \leq p \leq 0.6)$ if $p \sim \text{Beta}(6, 3)$:

```
## [1] 0.3141632
```

(Hmm. The first 4 digits are the same as $pi\ldots$)

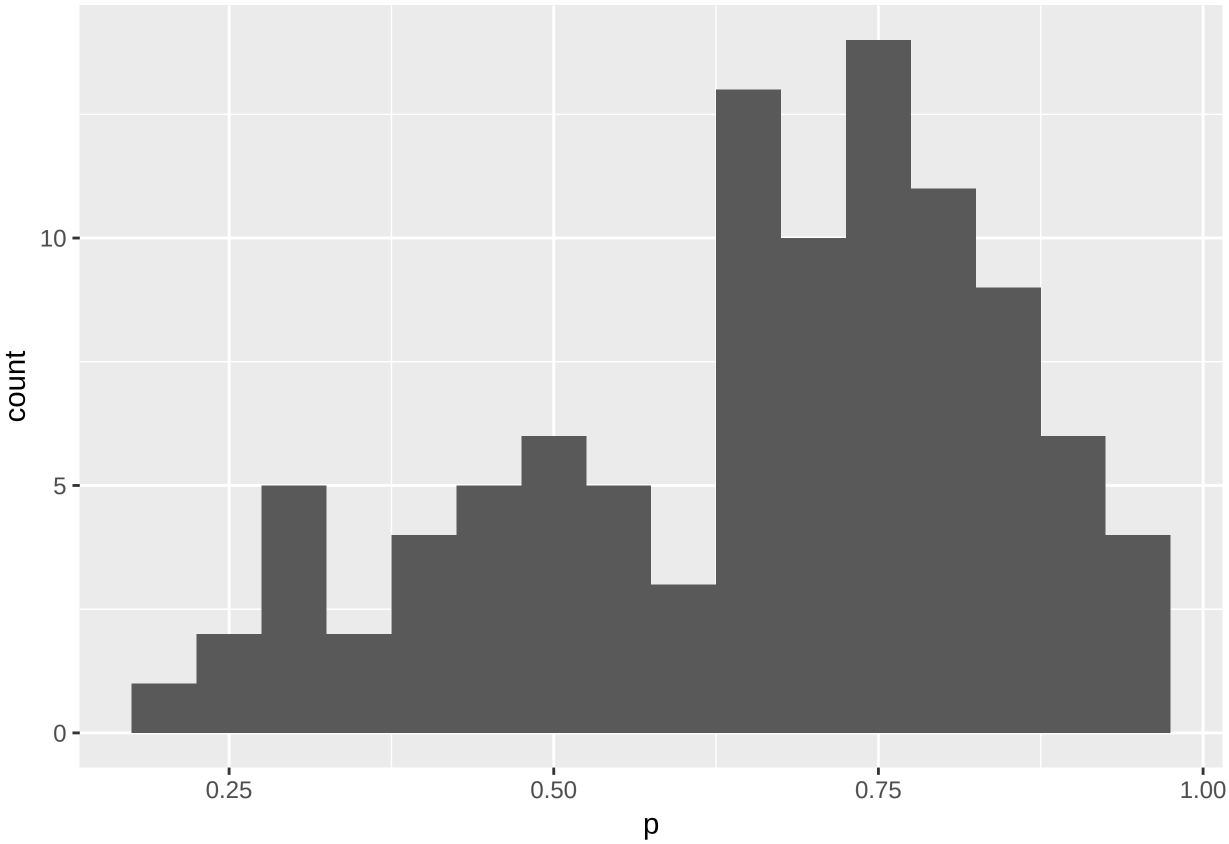
c) The quantile of the Beta$(10, 10)$ distribution at the probability values in the set $ {0.1, 0.5, 0.9, 1.5}$:

```
## Warning in qbeta(p, 10, 10): NaNs produced
```

```
##     p quantiles
## 1 0.1 0.3579299
## 2 0.5 0.5000000
## 3 0.9 0.6420701
## 4 1.5       NaN
```

d) A sample of 100 random values from Beta$(4, 2)$:

```
## Loading required package: ggplot2
```



Problem 4

We know that if $Z \sim \text{Beta}(a, b)$ then $E[X] = \frac{a}{a+b}$. So suppose $X = (p \mid Y = y) \sim \text{Beta}(a + y, b + n - y)$. Then

$$E[X] = \frac{a+y}{a+y+b+n-y} = \frac{a+y}{a+b+n}$$

Now since X is said to be a weighted average of the prior mean and the sample mean, suppose that $\exists \alpha \in (0, 1)$ such that $E[X] = \alpha E[p] + (1 - \alpha)\hat{p}$. We can expand the equation as follows:

$$\frac{a+y}{a+b+n} = \alpha \frac{a}{a+b} + (1-\alpha)\frac{y}{n} = \alpha\left(\frac{a}{a+b} - \frac{y}{n}\right) + \frac{y}{n}$$

Rearranging and then dividing by $\frac{a}{a+b} - \frac{y}{n}$ (when it is non-zero) on both sides we have that

$$\alpha = \frac{\frac{a+y}{a+b+n} - \frac{y}{n}}{\frac{a}{a+b} - \frac{y}{n}}$$

Expanding the fraction and then simplifying we have that

$$= \frac{\frac{na+ny-ay-by-ny}{n(a+b+n)}}{\frac{na-ay-by}{n(a+b)}} = \frac{\frac{na-ay-by}{n(a+b+n)}}{\frac{na-ay-by}{n(a+b)}} = \frac{a+b}{a+b+n}$$

Thus when $\frac{a}{a+b} \neq \frac{y}{n}$, $\alpha = \frac{a+b}{a+b+n}$. We can also show the sufficiency of α by reversing the argument. Thus,

$$E[X] = \frac{a+b}{a+b+n}E[p] + \frac{n}{a+b+n}\hat{p} = \frac{1}{1+\frac{n}{a+b}}E[p] + \frac{\frac{n}{a+b}}{1+\frac{n}{a+b}}\hat{p}$$

The posterior mean lies between the sample mean and the prior mean, thus integrating information from both model and experiment. Furthermore, the extent towards which it lies on either end depends how large n is with respect to $a+b$. The beta distribution reflects the confidence that among $a+b$ trials, on average, a will be successful. So when we conduct n trials, the posterior average is given by how the number of trials compared with the prior number of trials, $a+b$. We believe that more trials provide us more information about the value of p so if n is much smaller than $a+b$ then the posterior mean should be weighted more towards the prior mean since it relied on a higher number of trials. On the other hand, if n is much larger than $a+b$ then the posterior mean is weighted more towards the sample mean as a bigger sample reflects more information about the parameter. This can be seen from the fact that as $n \to \infty$, $\alpha \to 0$ so $1-\alpha \to 1$ and so the weight on the sample mean approaches 1. Finally, in the case where $\frac{a}{a+b} = \frac{y}{n}$, we see that the sample mean IS the prior mean so the sampling gave us no new information! That is why we get $E[X] = \frac{y}{n} = E[p]$.

Problem 5

Suppose the prior distribution $p \sim \text{Beta}(a,b)$. Then $\pi(p) = \frac{1}{B(a,b)}p^{a-1}(1-p)^{b-1}$ and likelihood $L(p) = \binom{n}{y}p^y(1-p)^{n-y}$. Then we can use Bayes' rule to get the posterior distribution $\pi(p \mid y)$:

$$\pi(p \mid y) = \frac{\pi(p)L(p)}{\int_0^1 \pi(t)L(t)\,dt}$$
$$= \frac{\frac{1}{B(a,b)}p^{a-1}(1-p)^{b-1}\binom{n}{y}p^y(1-p)^{n-y}}{\int_0^1 \frac{1}{B(a,b)}t^{a-1}(1-t)^{b-1}\binom{n}{y}t^y(1-t)^{n-y}\,dt}$$

Cancelling the $B(a,b)$ and the $\binom{n}{y}$ terms and then collecting the powers in the integral and the numerator, we get

$$\pi(p \mid y) = \frac{p^{a+y-1}(1-p)^{b+n-y-1}}{\int_0^1 t^{a+y-1}(1-t)^{b+n-y-1}\,dt}$$

Note that using the hint given during class, the integral at the bottom is just $B(a+y, b+n-y)$:

$$\pi(p \mid y) = \frac{1}{B(a+y, b+n-y)} p^{(a+y)-1}(1-p)^{(b+n-y)-1}$$

This is the density function for $\text{Beta}(a+y, b+n-y)$. Since we can identify distributions with their density functions, we have that $p \mid Y = y \sim \text{Beta}(a+y, b+n-y)$