# Bayesian Statistics- HW2

魏上傑

2023-04-17

1.

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.

$p = \{0.3, 0.4, 0.5, 0.6, 0.7, 0.8\}$

$\pi_{owner}(p) = (0.125, 0.125, 0.250, 0.250, 0.125, 0.125)$

```
bayes_table <- data.frame(p=seq(0.3,0.8,by=0.1),
                          Prior=c(0.125, 0.125, 0.250,
                                  0.250, 0.125, 0.125))
ggplot(data=bayes_table, aes(x=p, y=Prior))+
  geom_bar(stat = "identity", fill=crcblue, width=0.06)
```
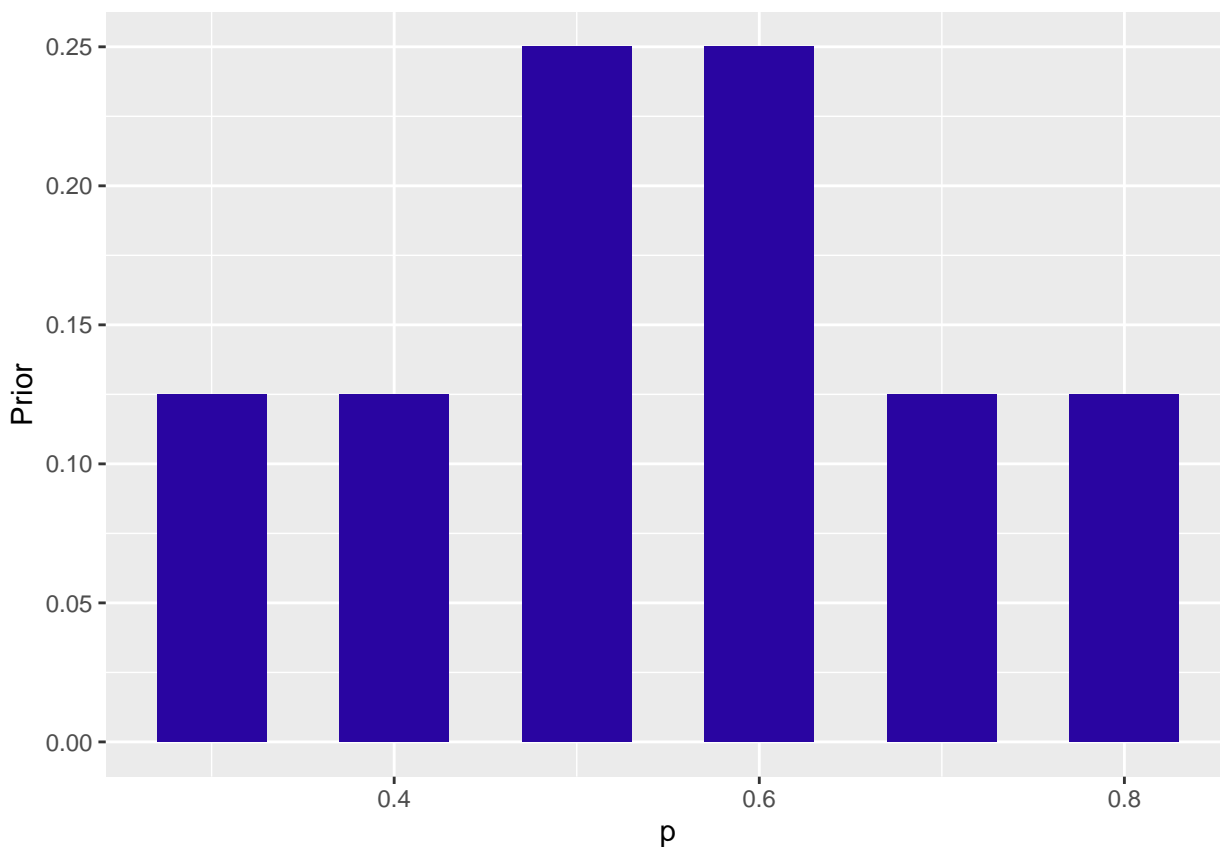
```r
bayes_table$Likelihood <- dbinom(10, size=30,
                                 prob = bayes_table$p)
bayes_table
```

```
##     p Prior   Likelihood
## 1 0.3 0.125 1.415617e-01
## 2 0.4 0.125 1.151854e-01
## 3 0.5 0.250 2.798160e-02
## 4 0.6 0.250 1.997491e-03
## 5 0.7 0.125 2.959225e-05
## 6 0.8 0.125 3.382768e-08
```

```r
bayes_table <- bayesian_crank(bayes_table)
bayes_table
```

```
##     p Prior   Likelihood      Product    Posterior
## 1 0.3 0.125 1.415617e-01 1.769521e-02 4.469407e-01
## 2 0.4 0.125 1.151854e-01 1.439817e-02 3.636649e-01
## 3 0.5 0.250 2.798160e-02 6.995400e-03 1.766878e-01
## 4 0.6 0.250 1.997491e-03 4.993727e-04 1.261301e-02
## 5 0.7 0.125 2.959225e-05 3.699031e-06 9.342907e-05
## 6 0.8 0.125 3.382768e-08 4.228460e-09 1.068012e-07
```

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

```
sum(bayes_table$Posterior[bayes_table$p==0.3])
```

```
## [1] 0.4469407
```

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

```
sum(bayes_table$Posterior[bayes_table$p>0.5])
```

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

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

```
sum(bayes_table$Posterior[bayes_table$p>0.2 & bayes_table$p<0.4])
```

```
## [1] 0.4469407
```

  2.

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.

Beta(1, 1) represents the prior belief that each value of p is equally likely from zero to one.

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

Beta(0.5, 1) represents the prior belief that the extreme value $p = 0$ is the most probable, which indicates that the owner thinks the students' preference of dining out on Friday is very weak.

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

Beta(4, 2) represents the prior belief that the values of p around 0.75 are more probable.

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

Beta(4, 1) represents the prior belief that the values of p becomes more probable as the values of p increase.

This is different from Beta(4, 2), where there's a peak around p=0.75.

For Beta(4, 1), the owner thinks the students' preference of dining out on Friday is strong, even for extreme value p=1.

For Beta(4, 2), the owner thinks the students' preference of dining out on Friday is strong to some extent, but not for extreme value p=1.

3.

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}.

```
dbeta(0.1, 0.5, 0.5)
```

```
## [1] 1.061033
```

```
dbeta(0.5, 0.5, 0.5)
```

```
## [1] 0.6366198
```

```
dbeta(0.9, 0.5, 0.5)
```
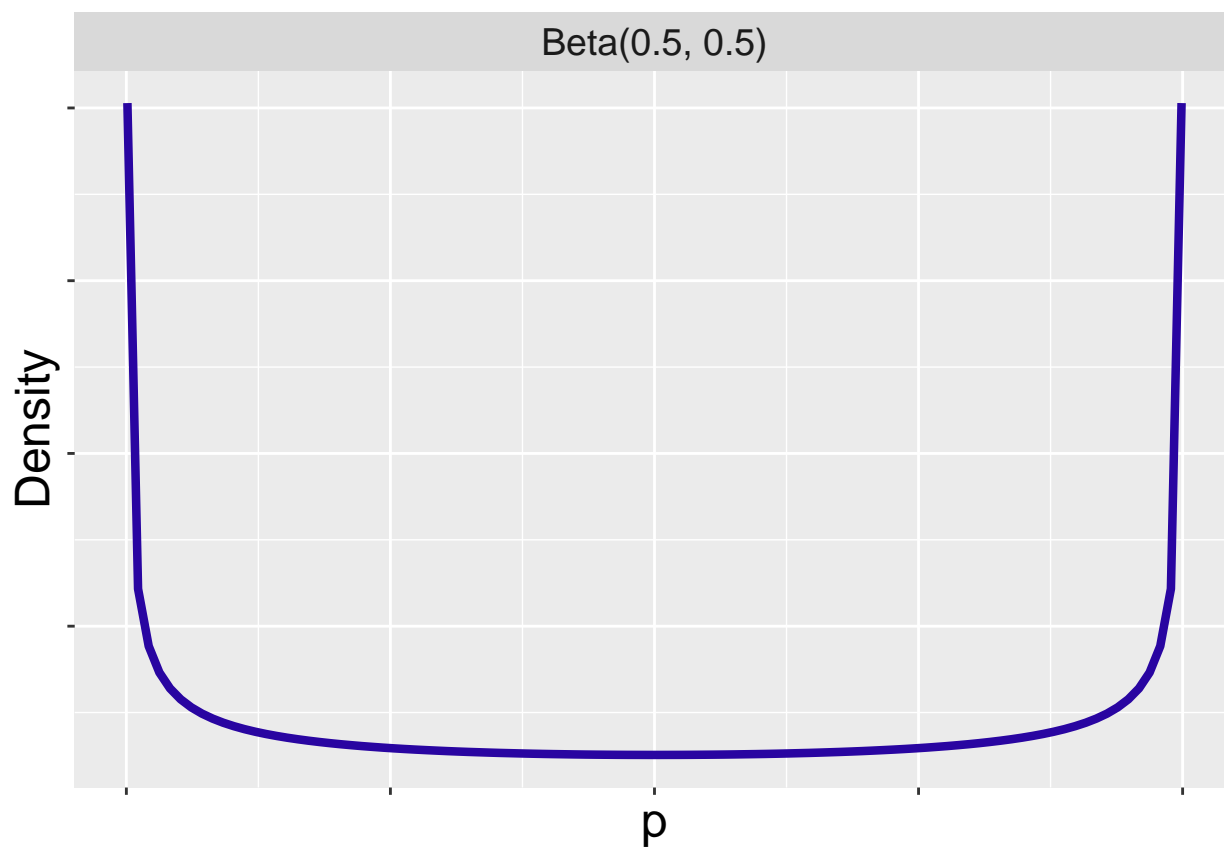
```
## [1] 1.061033
```

```
dbeta(1.5, 0.5, 0.5)
```

```
## [1] 0
```

```
p <- seq(.001, .999, length.out = 100)
BETA <- data.frame(p = p, Density = dbeta(p, 0.5, 0.5), Type= "Beta(0.5, 0.5)")

ggplot(BETA, aes(p, Density))+
  geom_line(color = crcblue, linewidth=1.5) +
  facet_wrap(~ Type, scales = "free") +
  increasefont() +
  theme(axis.text= element_blank())
```
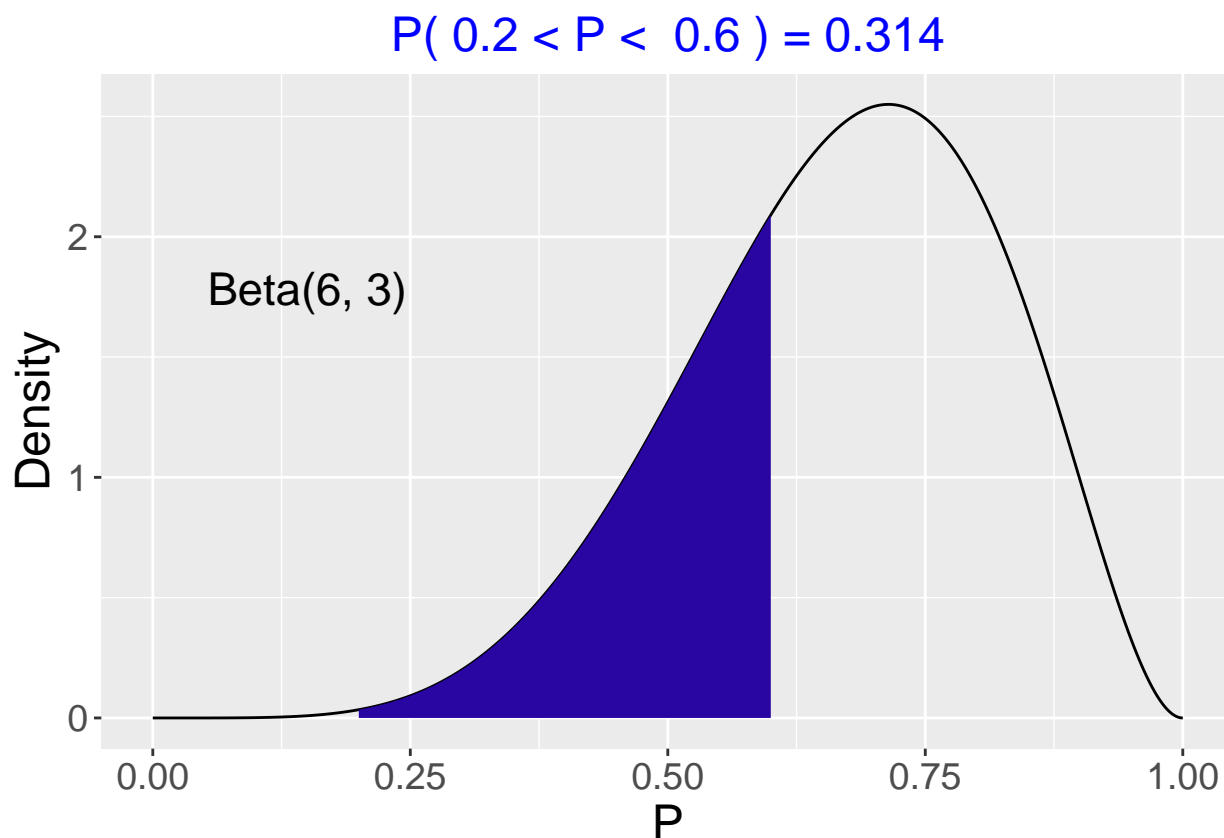
(b) The probability P $(0.2 \leq p \leq 0.6)$ if $p \sim Beta(6, 3)$.

```
pbeta(0.6, 6, 3, lower.tail = TRUE) - pbeta(0.2, 6, 3, lower.tail = TRUE)
```

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

```
beta_area(lo=0.2, hi=0.6, shape_par = c(6, 3), Color = crcblue) +
  theme(text = element_text(size=18))
```

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

```
qbeta(0.1, 10, 10, lower.tail = TRUE)
```

```
## [1] 0.3579299
```

```
qbeta(0.5, 10, 10, lower.tail = TRUE)
```

```
## [1] 0.5
```

```
qbeta(0.8, 10, 10, lower.tail = TRUE)
```
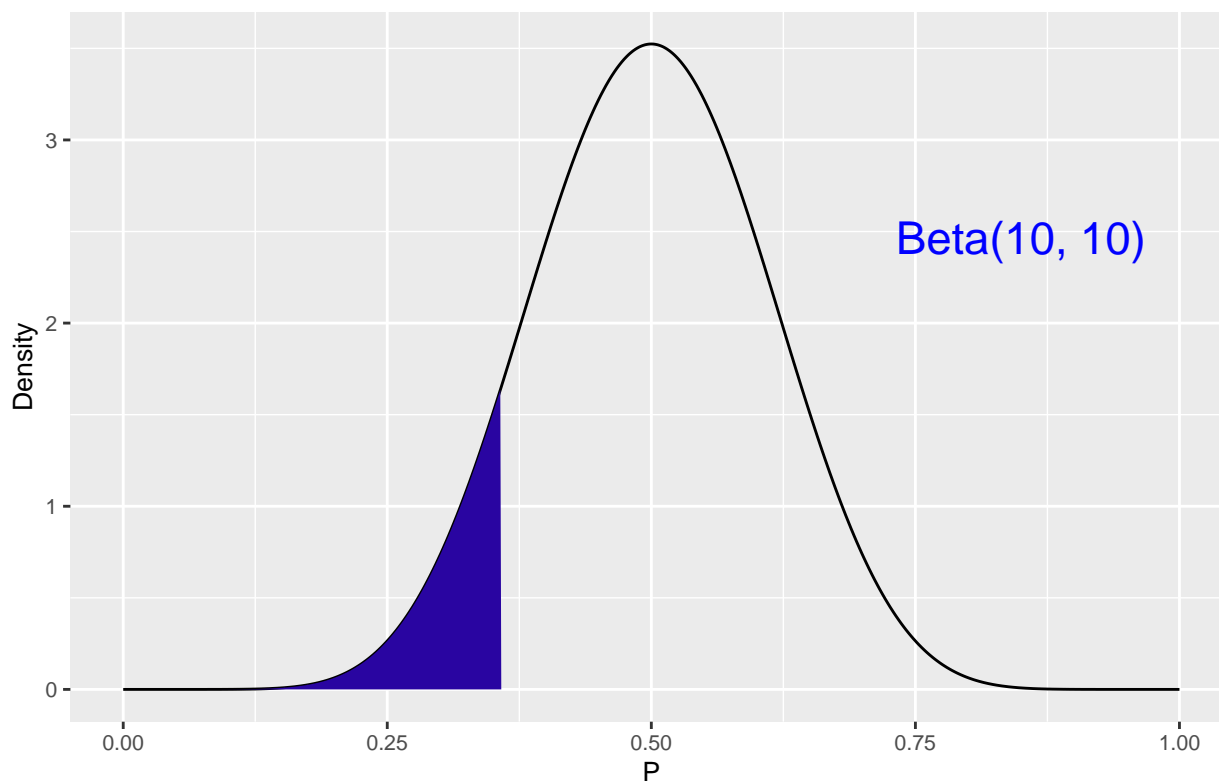
```
## [1] 0.5944172
```

```
qbeta(1, 10, 10, lower.tail = TRUE)
```
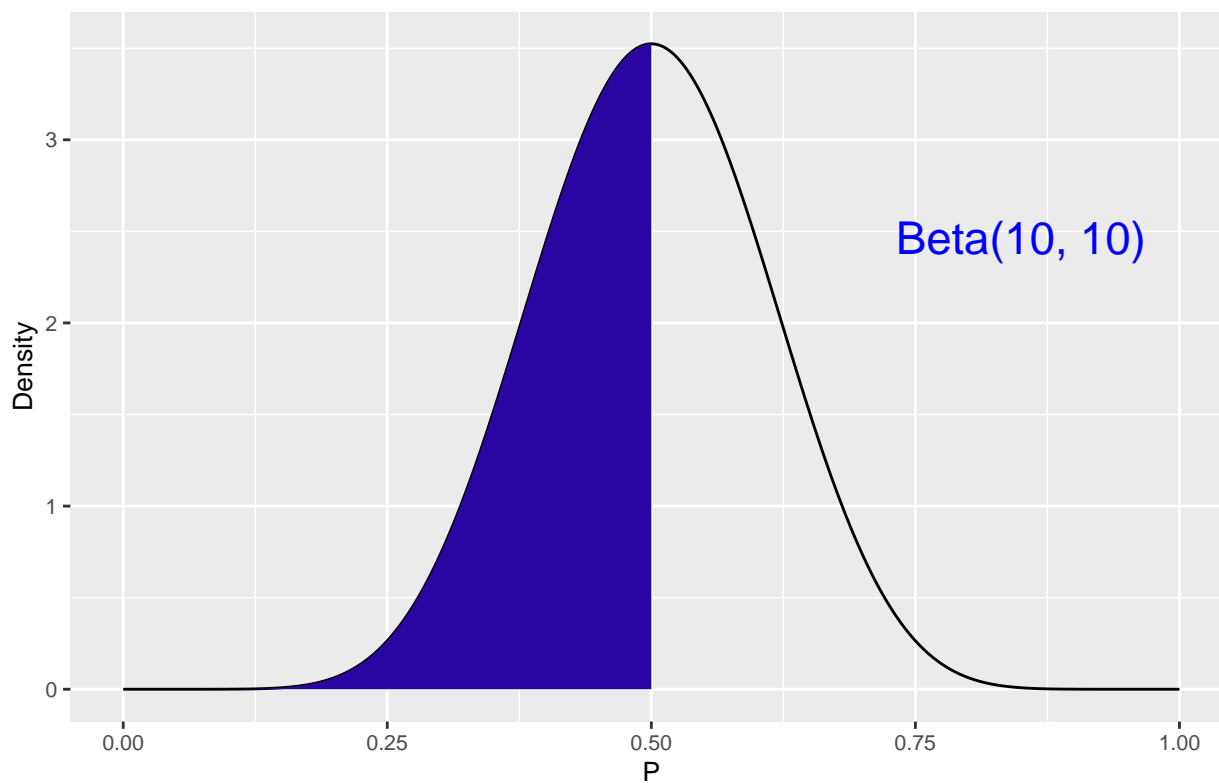
```
## [1] 1
```

```
beta_quantile(0.1, c(10, 10), Color = crcblue) +
  theme(text = element_text(size=10))
```
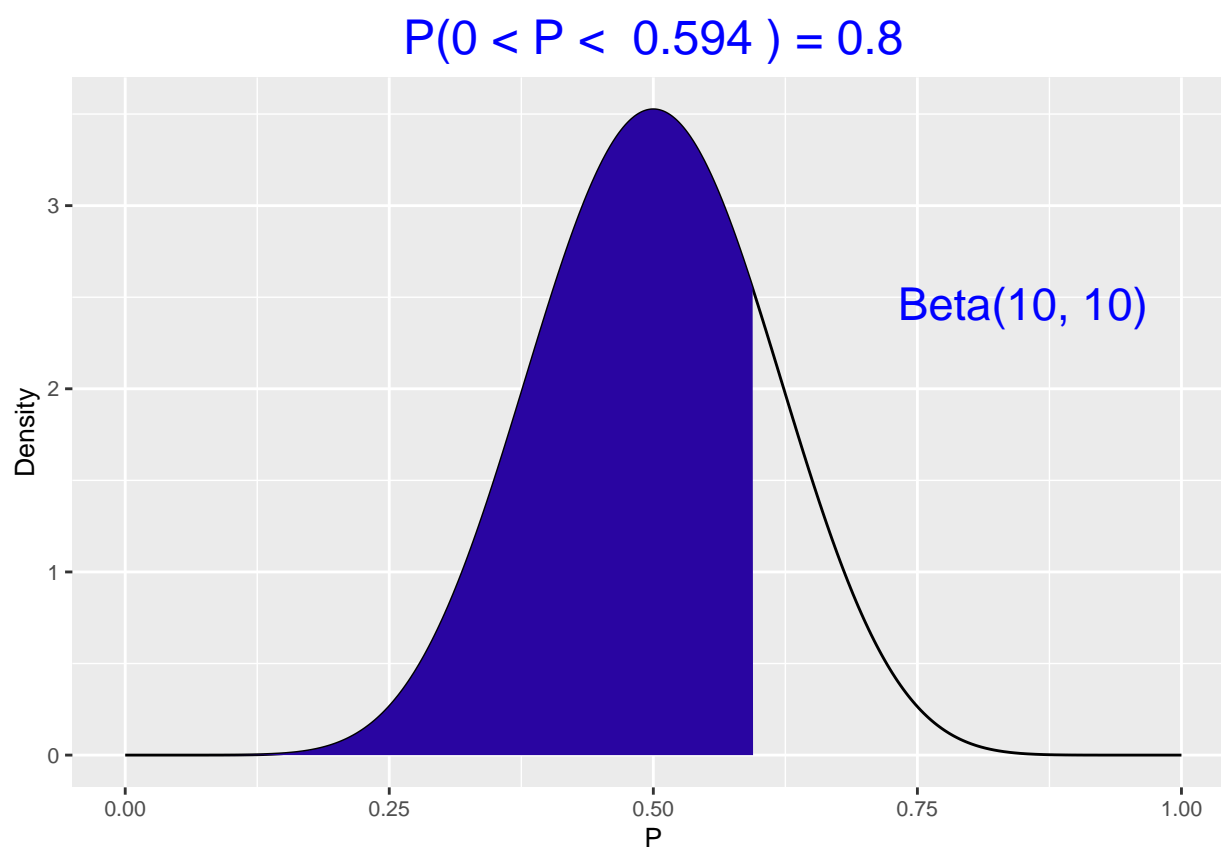
```
beta_quantile(0.5, c(10, 10), Color = crcblue) +
  theme(text = element_text(size=10))
```

```
beta_quantile(0.8, c(10, 10), Color = crcblue) +
    theme(text = element_text(size=10))
```



```
beta_quantile(1, c(10, 10), Color = crcblue) +
    theme(text = element_text(size=10))
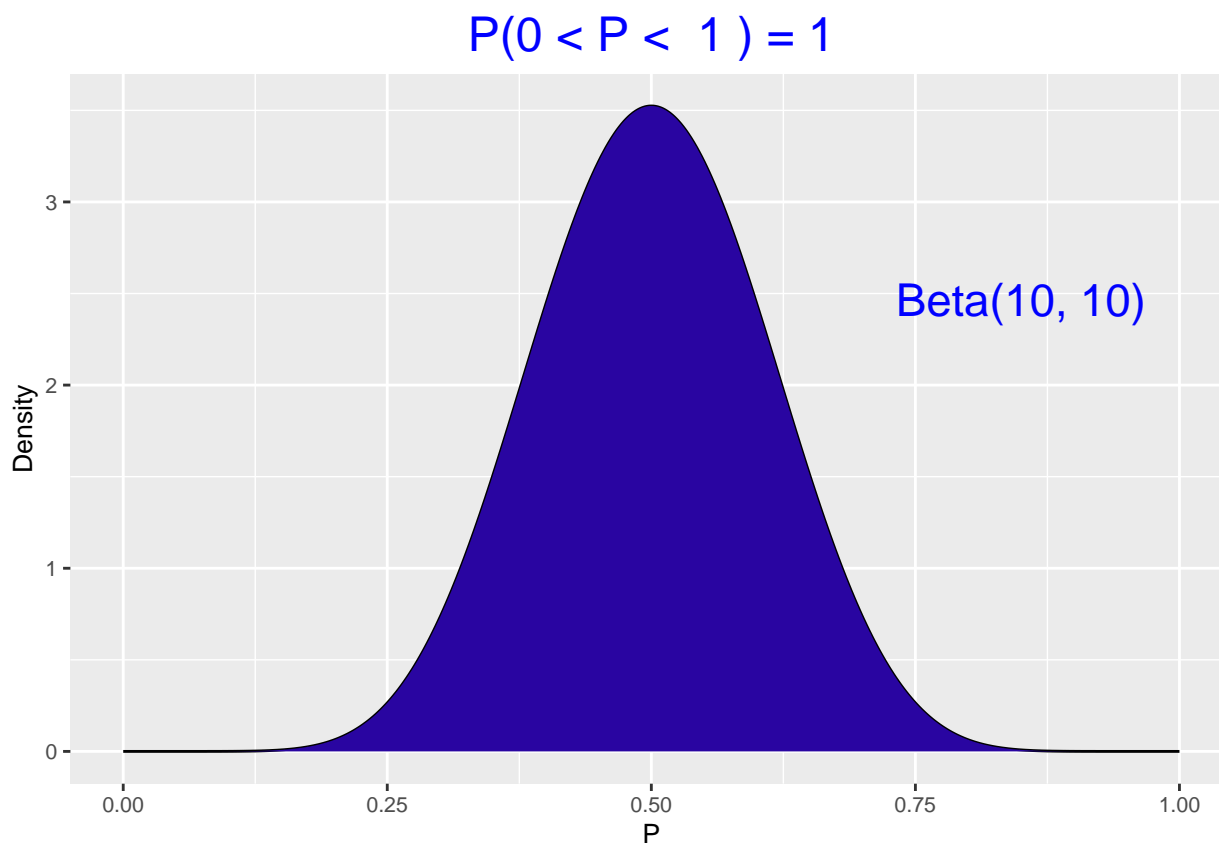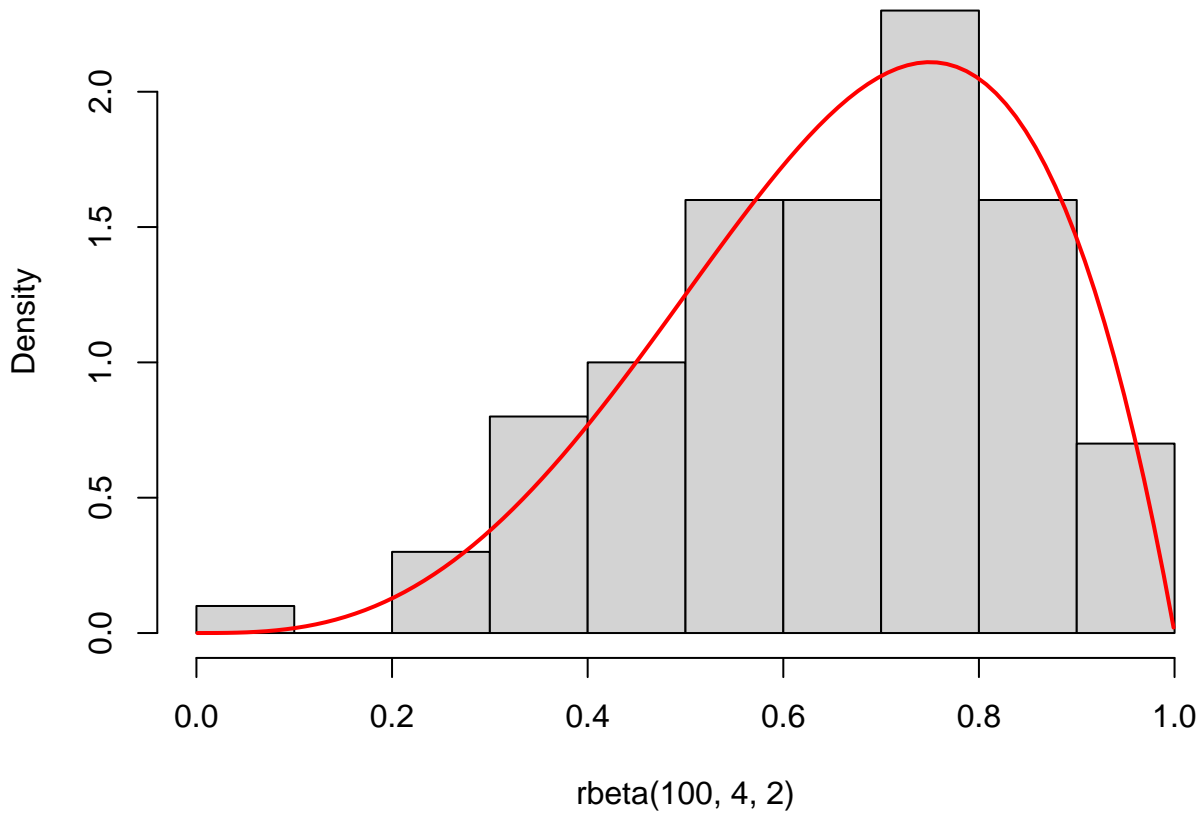```

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

```
set.seed(1111)
p <- seq(.001, .999, length.out = 100)
par(mai=c(0.9, 0.9, 0.1, 0.1))
hist(rbeta(100, 4, 2), probability = TRUE, main = NULL)
lines(p, dbeta(p, 4, 2), col="red", lwd=2)
```

rbeta(100, 4, 2)

4.

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 $p|Y = y \sim Beta(a + y, b + n - y)$ 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.

By the fact that the mean of a Beta(a, b) random variable following is $\frac{a}{a+b}$, the mean of $p|Y = y$ is
$\frac{a+y}{a+y+b+n-y} = \frac{a+y}{a+b+n} = \frac{a}{a+b+n} + \frac{y}{a+b+n} = w_1 \frac{a}{a+b} + w_2 \frac{y}{n}$

This implies $w_1 = \frac{a+b}{a+b+n}, w_2 = \frac{n}{a+b+n}$

We now have

$$\frac{a+b}{a+b+n}\frac{a}{a+b} + \frac{n}{a+b+n}\frac{y}{n} \tag{1}$$

Hence, if one has strong prior (large a and/or b), then the posterior will be heavily influenced by prior. On the other hand, if one has strong data (large n), then the posterior will be heavily influenced by data.

Also, note that the prior mean is $\frac{a}{a+b}$, thus, equation (1) shows collected data help sharpen one's belief from prior to posterior, instead of directly determining what one's belief should be after seeing the data.

5.

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.

Derivations:

$$p \sim Beta(a, b) \implies \pi(p) = \frac{1}{B(a,b)} p^{a-1}(1-p)^{b-1}$$

$$y|p \sim Bin(n, p) \implies L(p) = \binom{n}{y} p^y (1-p)^{n-y}$$

Want to show: $p|y \sim Beta(a + y, b + n - y)$

$$\pi(p|y) = \frac{\pi(p)L(p)}{\int \pi(\tilde{p})L(\tilde{p})d\tilde{p}}$$

$$= \frac{\frac{1}{B(a,b)} p^{a-1}(1-p)^{b-1} \binom{n}{y} p^y (1-p)^{n-y}}{\int \frac{1}{B(a,b)} \tilde{p}^{a-1}(1-\tilde{p})^{b-1} \binom{n}{y} \tilde{p}^y (1-\tilde{p})^{n-y} d\tilde{p}}$$

$$= \frac{p^{a+y-1}(1-p)^{b+n-y-1}}{\int \tilde{p}^{a+y-1}(1-\tilde{p})^{b+n-y-1} d\tilde{p}}$$

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

The last equation comes about by the fact that:

$$\int \pi(p)dp = 1 \implies \int \frac{1}{B(a+y,b+n-y)} \tilde{p}^{a+y-1}(1-\tilde{p}^{b+n-y-1})d\tilde{p} = 1$$

Hence, $\pi(p|y) \sim Beta(a + y, b + n - y)$