ST4234_T1

Bian Ce

2023-01-27

Q1(b) [Hoff, Exercise 3.1]

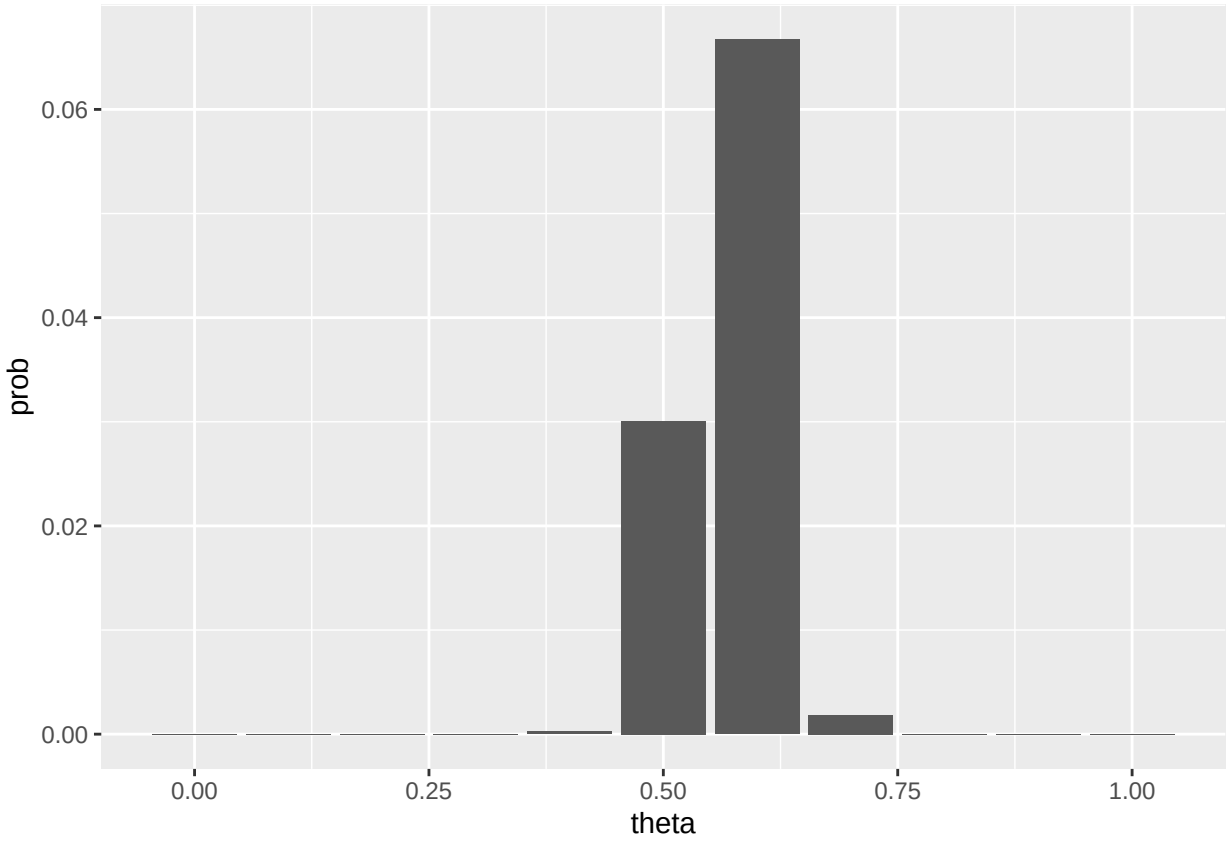
```
Y = 57
n = 100
theta = seq(from = 0, to = 1, by = 0.1)

probs = dbinom(x=Y, size=n, p=theta)

df.1b = data.frame(theta = theta, prob = probs); df.1b
```

```
##    theta         prob
## 1    0.0 0.000000e+00
## 2    0.1 4.107157e-31
## 3    0.2 3.738459e-16
## 4    0.3 1.306895e-08
## 5    0.4 2.285792e-04
## 6    0.5 3.006864e-02
## 7    0.6 6.672895e-02
## 8    0.7 1.853172e-03
## 9    0.8 1.003535e-07
## 10   0.9 9.395858e-18
## 11   1.0 0.000000e+00
```

```
ggplot(df.1b,aes(x=theta, y=prob)) + geom_bar(stat = 'identity')
```



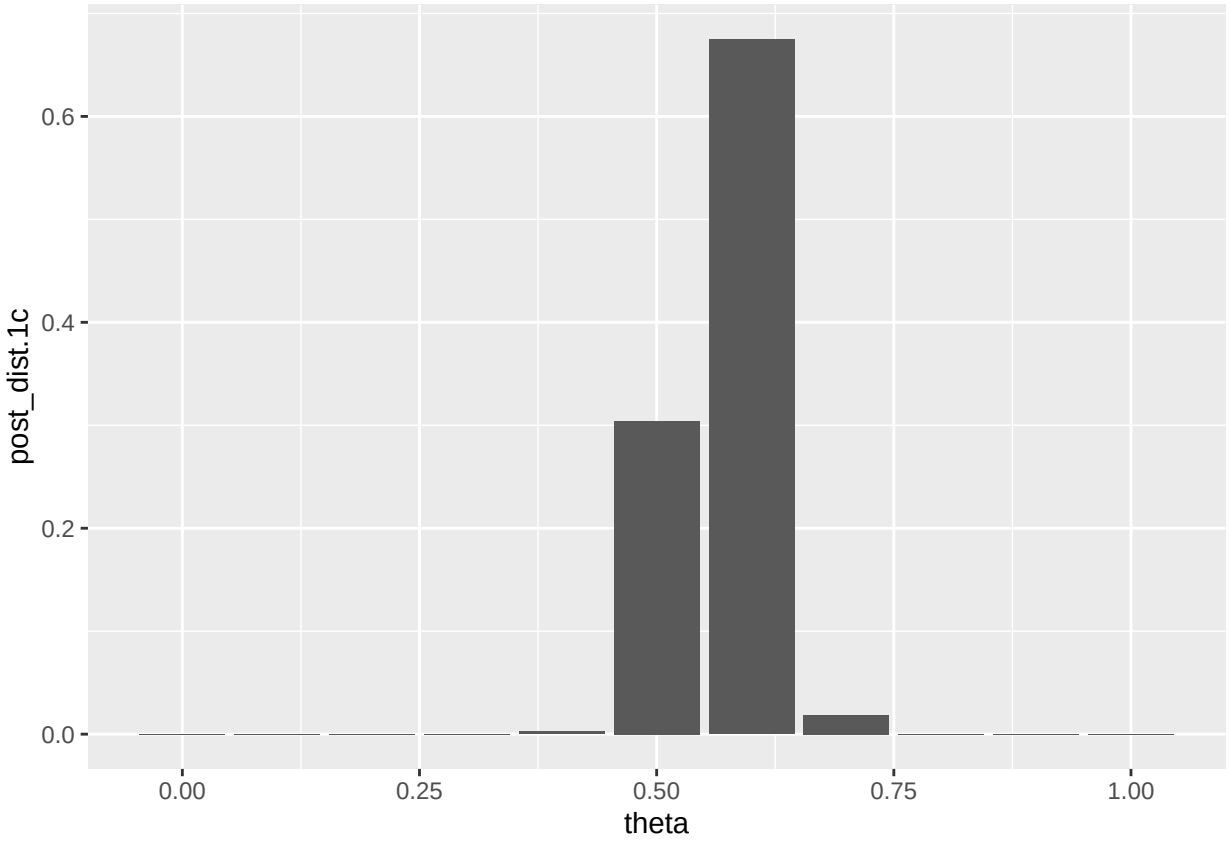
Q1(c)

```
denom <- sum(dbinom(Y, n, theta))
post.1c <- dbinom(Y, n, theta) / denom

df.1c <- data.frame(theta=theta, post_dist.1c = post.1c); df.1c
```

```
##    theta post_dist.1c
## 1    0.0 0.000000e+00
## 2    0.1 4.153701e-30
## 3    0.2 3.780824e-15
## 4    0.3 1.321705e-07
## 5    0.4 2.311695e-03
## 6    0.5 3.040939e-01
## 7    0.6 6.748515e-01
## 8    0.7 1.874172e-02
## 9    0.8 1.014907e-06
## 10   0.9 9.502335e-17
## 11   1.0 0.000000e+00
```

```
ggplot(df.1c, aes(x=theta, y=post_dist.1c)) + geom_bar(stat = 'identity')
```




```
sum(df.1c$post_dist)
```

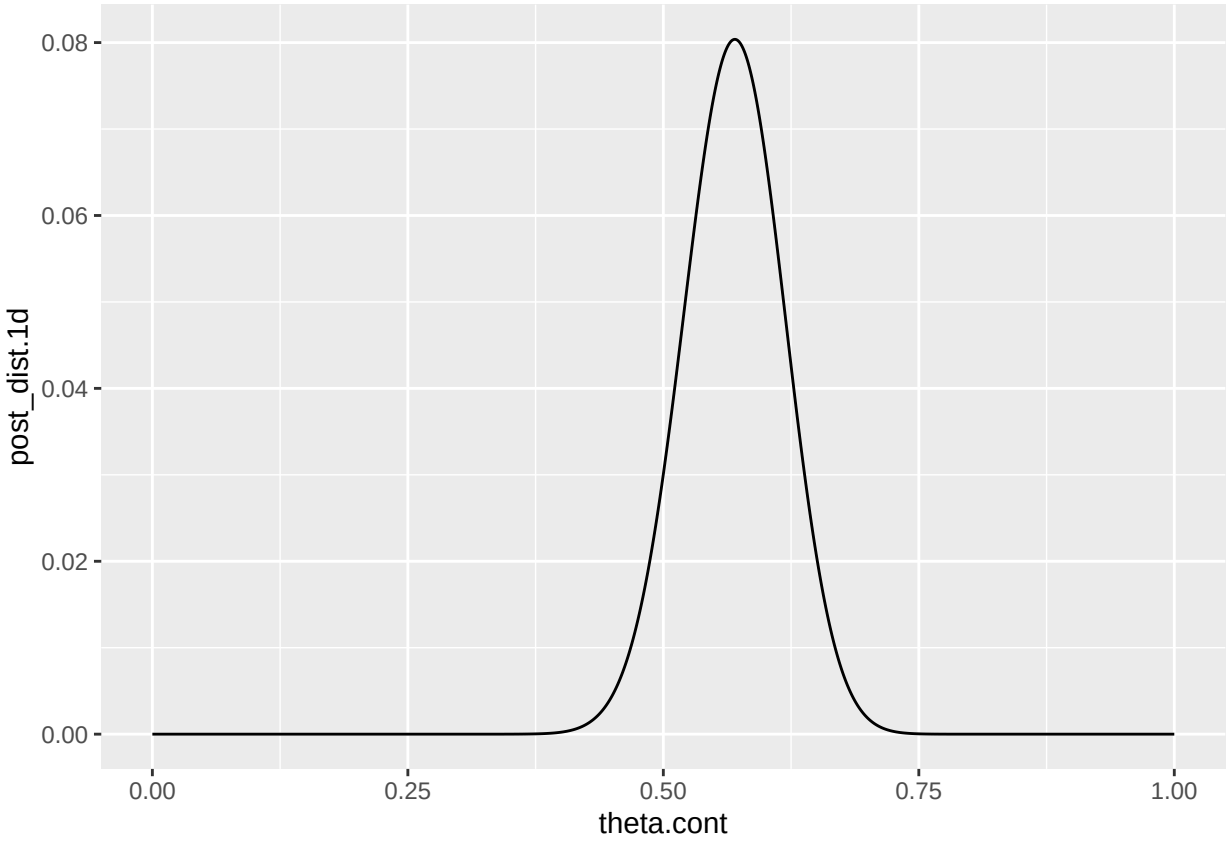
```
## [1] 1
```

Q1(d)

```
theta.cont <- seq(0, 1, by=0.0001)
post.1d <- dbinom(Y, n, theta.cont)

df.1d <- data.frame(theta.cont=theta.cont, post_dist.1d=post.1d)

ggplot(df.1d, aes(x=theta.cont, y=post_dist.1d)) + geom_line()
```

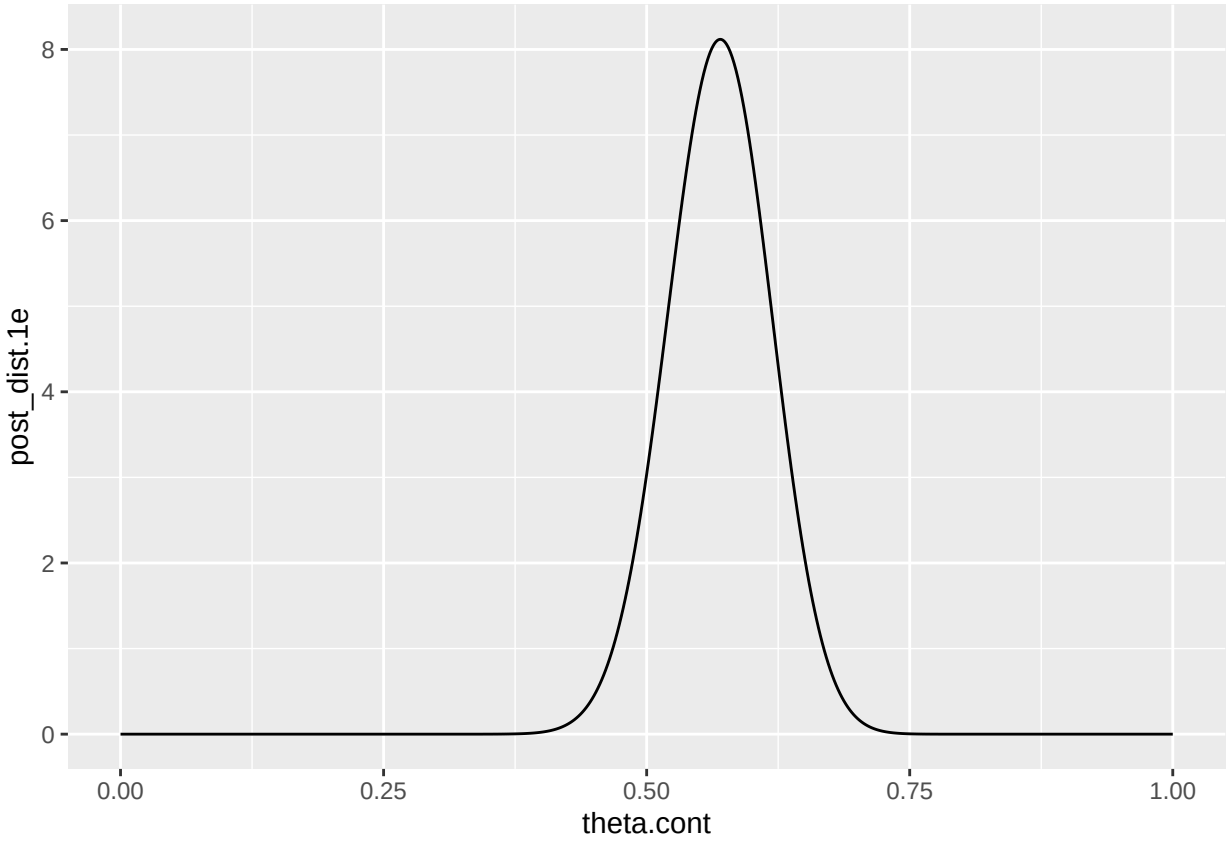


Q1(e)

```
post.1e <- dbeta(theta.cont, 58,44)

df.1e <- data.frame(theta.cont=theta.cont, post_dist.1e=post.1e)

ggplot(df.1e, aes(x=theta.cont, y=post_dist.1e)) + geom_line()
```



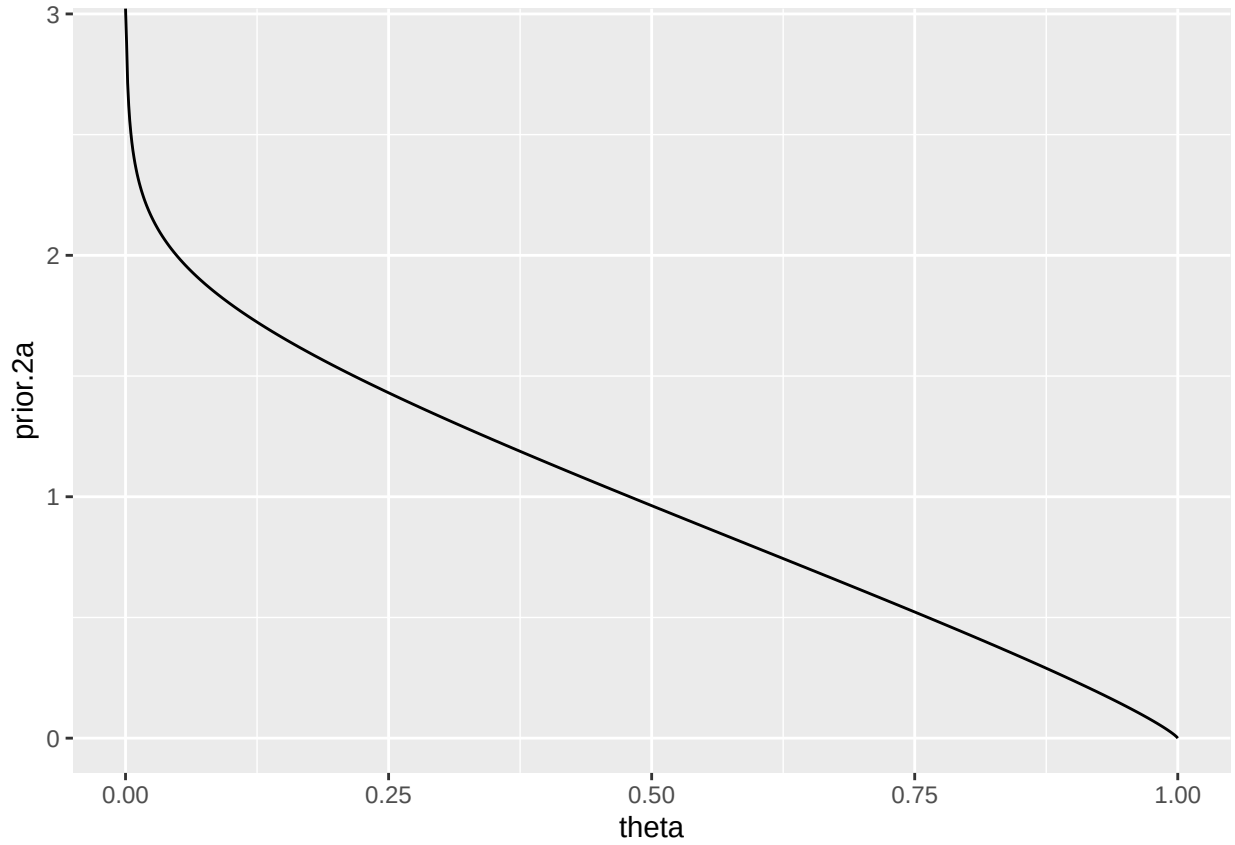
Q2(a)

```
beta_mean <- function(a, b){
  return (a/(a+b))
}

beta_var <- function(a, b) {
  return (a*b / ((a+b)^2 * (a+b+1)))
}

rm(theta)
theta <- seq(0, 1, by=0.001)
a.2a = 11/12; b.2a = 11/6
prior.2a <- dbeta(theta, a.2a, b.2a)

df.2a <- data.frame(theta=theta, prior.2a=prior.2a);
ggplot(df.2a, aes(x=theta, y=prior.2a)) + geom_line()
```




```
prior.2a.sd = sqrt(beta_var(a.2a, b.2a))
prior.2a.sd
```

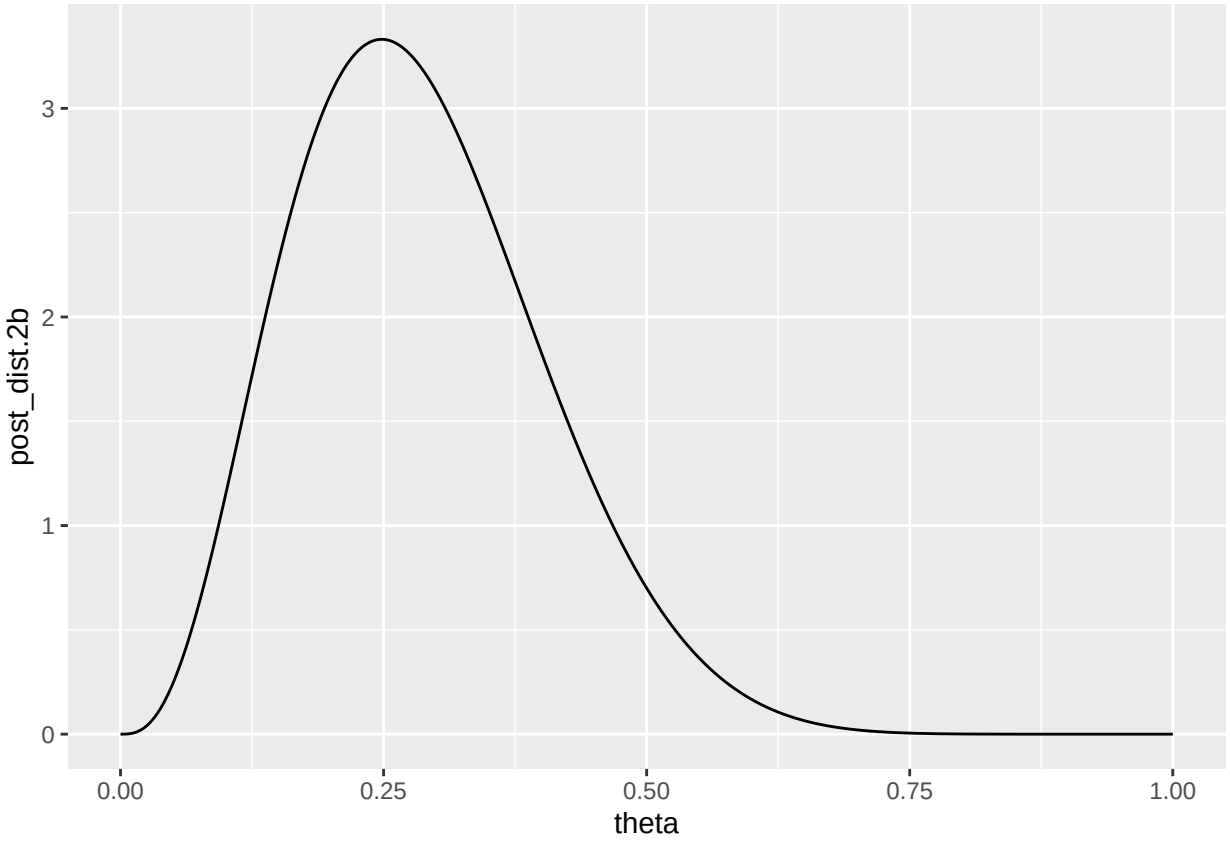
```
## [1] 0.2434322
```

Q2(b)

```
a.2b = 47/12; b.2b = 59/6
post.2b <- dbeta(theta, a.2b, b.2b)

df.2b <- data.frame(theta=theta, post_dist.2b=post.2b)

ggplot(df.2b, aes(x=theta, y=post_dist.2b)) + geom_line()
```




```
mean.2b = beta_mean(a.2b, b.2b); mean.2b
```

```
## [1] 0.2848485
```

```
median.2b = qbeta(0.5, a.2b, b.2b); median.2b
```

```
## [1] 0.2741738
```

```
mode.2b = (a.2b-1)/(a.2b+b.2b-2); mode.2b
```

```
## [1] 0.248227
```

Q2(c)

```
require(TeachingDemos)

hpd.50 <- hpd(qbeta, shape1=a.2b, shape2=b.2b, conf=0.50); hpd.50
```

```
## [1] 0.1738371 0.3342303
```

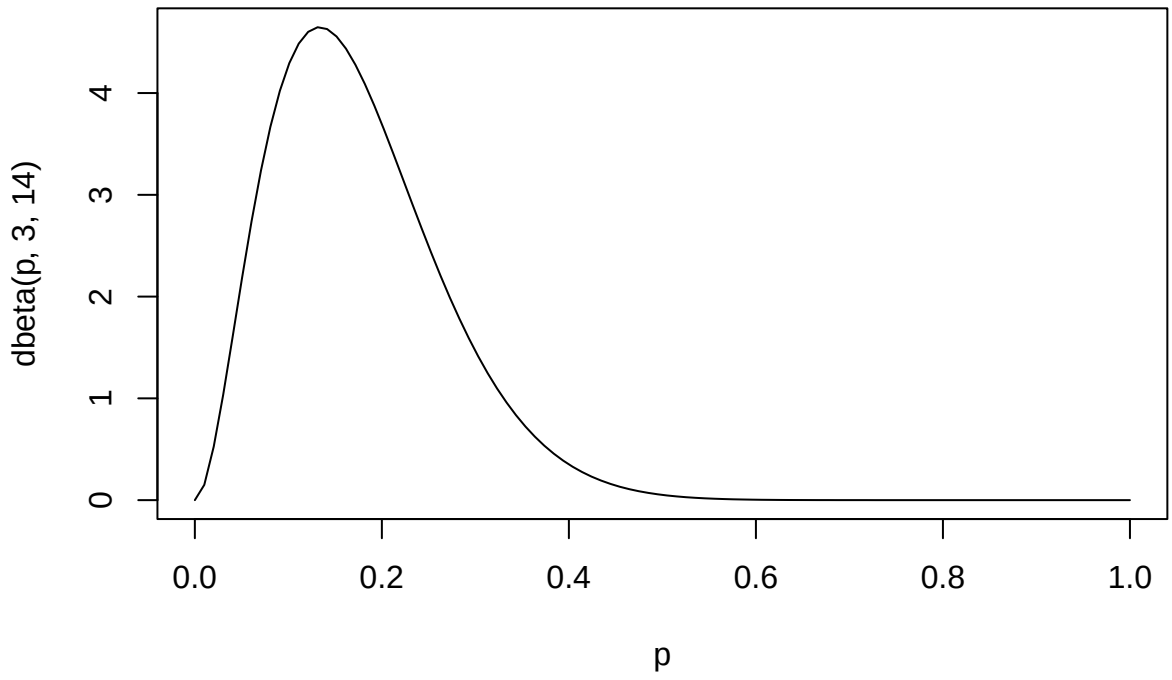
```
quant.50 <- qbeta(c(0.25, 0.75), a.2b, b.2b); quant.50
```

```
## [1] 0.1975017 0.3611453
```

Q3(a) [Hoff, Exercise 3.7]

```
#define range
p = seq(0,1, length=100)

#create plot of Beta distribution with shape parameters 2 and 10
plot(p, dbeta(p, 3, 14), type='l')
```

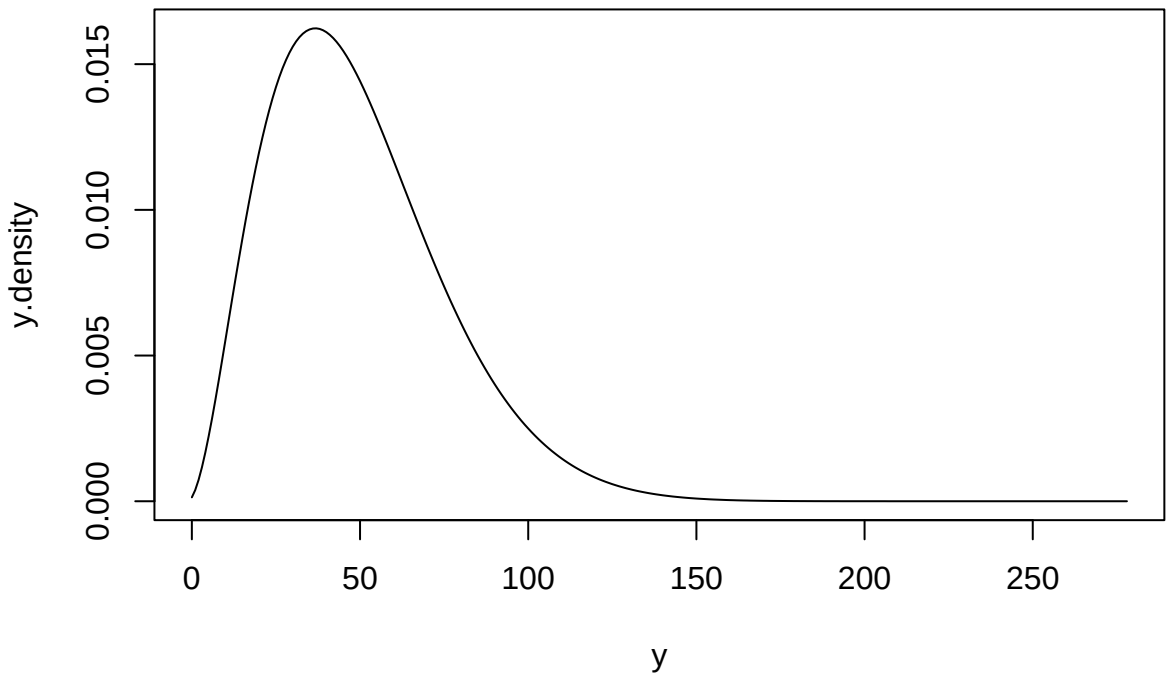


Q3(c)

```
rm(n)

n = 278
Pr <- function(y) {
  return (choose(n, y) * beta(y+3, 292-y) / beta(3,14))
}

y <- seq(0, n, by = 1)
y.density <- sapply(y, Pr)
plot(y, y.density, type='l')
```




```
# Obtain the mean and s.d.
y.mean <- sum(y*y.density)
y.mean # 49.05882
```

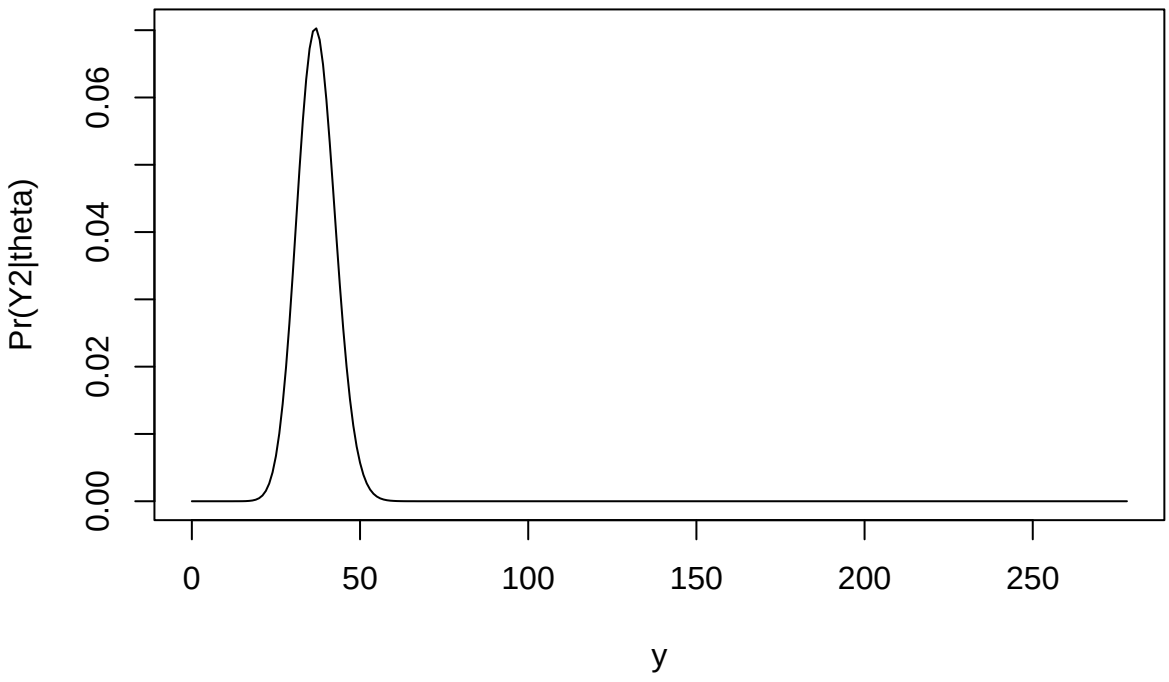
```
## [1] 49.05882
```

```
y.sd <- sqrt(sum(y.density * (y - y.mean)^2))
y.sd # 25.73196
```

```
## [1] 25.73196
```

Q3(d)

```
theta_hat = 2/15
y.mle.density <- dbinom(y, n, theta_hat)
plot(y, y.mle.density, xlab="y", ylab="Pr(Y2|theta)", type = 'l')
```




```
bino_mean <- function(theta,n) {n*theta}
bino_sd <- function(theta,n) {sqrt(n*theta*(1-theta))}
y.mle.mean <- bino_mean(theta_hat ,n)
y.mle.mean # 37.06667
```

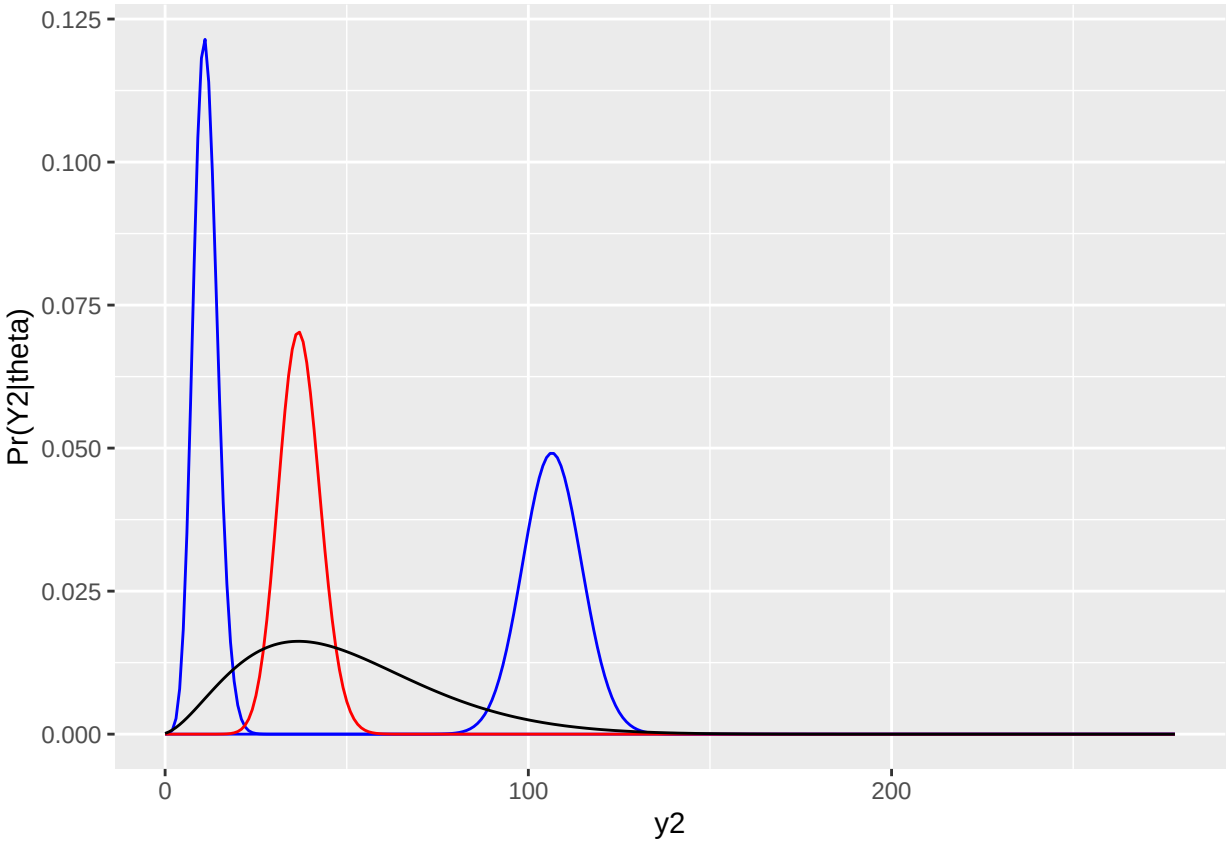
```
## [1] 37.06667
```

```
y.mle.sd <- bino_sd(theta_hat, n)
y.mle.sd # 5.667843
```

```
## [1] 5.667843
```

```
# Calculating the 95% c.i. for theta
theta.ci <- qbeta(c(0.025, 0.975), 3, 14)
df <- data.frame(y, dbinom(y, n, theta.ci[1]), dbinom(y, n, theta.ci[2]))

ggplot(df, aes(y)) +
  geom_line(aes(y = dbinom(y, n, theta.ci[1])), colour="#0000FF") +
  geom_line(aes(y = dbinom(y, n, theta.ci[2])), colour="#0000FF") +
  geom_line(aes(y = y.mle.density), colour="#FF0000") +
  geom_line(aes(y = y.density), colour="#000000") +
  xlab("y2") + ylab("Pr(Y2|theta)")
```




```
y.left.mean <- bino_mean(theta.ci[1], n)
y.left.sd <- bino_sd(theta.ci[1], n)
y.left.mean # 11.2517
```

```
## [1] 11.2517
```

```
y.left.sd # 3.285772
```

```
## [1] 3.285772
```

```
y.right.mean <- bino_mean(theta.ci[2], n)
y.right.sd <- bino_sd(theta.ci[2], n)
y.right.mean # 106.6064
```

```
## [1] 106.6064
```

```
y.right.sd # 8.107119
```

```
## [1] 8.107119
```