Bayesian Stats Assignment-2

Mayank Kumar Raunak(mraunak@iu.edu)

October 2, 2019

R Markdown

3.1 Sample survey Suppose we are going to sample 100 individuals from a county (of size much larger than 100) and ask each sampled person whether they support policy Z or not. Let Yi = 1 if person i in the sample supports the policy, and Yi = 0 otherwise.

a) Assume Y1, . . . , Y100 are, conditional on θ , i.i.d. binary random variables with expectation θ. Write down the joint distribution of $Pr(Y1 = y1, \dots, Y100 = y100|\theta)$ in a compact form. Also write down the form of $Pr(\sum PY_i = y|\theta)$.

solution: The joint probability distribution of $Pr(Y1 = y1, \dots, Y100 = y100|\theta)$ in compact form is:

$p(Y1 = y1, \dots, Y100 = y100|\theta) = \theta^{\sum_{i=1}^{100} y_i}(1-\theta)^{100 \ \ \sum_{i=1}^{100} y_i}$

$Pr(\sum PY_i = y|\theta)$ = $(100_{c_y})\theta^y(1-\theta)^{100-y}$

b) For the moment, suppose you believed that θ = {0.0, 0.1, . . . , 0.9, 1.0}.Given that the results of the survey were $\sum_{i=1}^{100} Y_i = 57$, compute $Pr(PY_i = 57|\theta)$ for each of these 11 values of θ and plot these probabilities as a function of θ.

Solution: $Pr(PY_i = 57|\theta = 0.0)$ = dbinom(57,100,0) = 0

$Pr(PY_i = 57|\theta = 0.1)$ = dbinom(57,100,0.1) = 4.107157e-31

$Pr(PY_i = 57|\theta = 0.2)$ = dbinom(57,100,0.2) = 3.738459e-16

$Pr(PY_i = 57|\theta = 0.3)$ = dbinom(57,100,0.3) = 1.306895e-08

$Pr(PY_i = 57|\theta = 0.4)$ = dbinom(57,100,0.4) = 0.0002285792

$Pr(PY_i = 57|\theta = 0.5)$ = dbinom(57,100,0.5) = 0.03006864

$Pr(PY_i = 57|\theta = 0.6)$ = dbinom(57,100,0.6) = 0.06672895

$Pr(PY_i = 57|\theta = 0.7)$ = dbinom(57,100,0.7) = 0.001853172

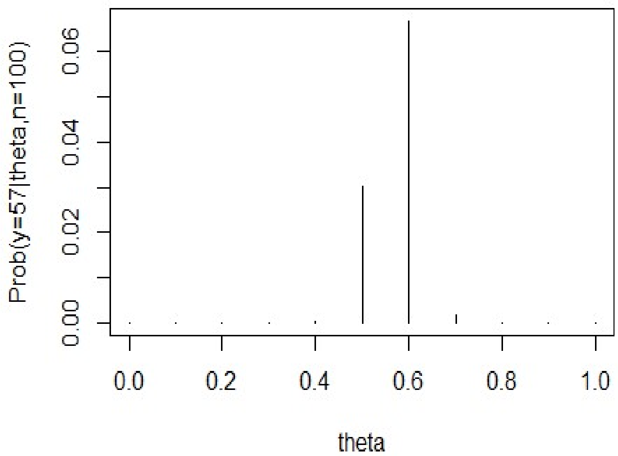
$Pr(PY_i = 57|\theta = 0.8)$ = dbinom(57,100,0.8) = 1.003535e-07

$Pr(PY_i = 57|\theta = 0.9)$ = dbinom(57,100,0.9) = 9.395858e-18

$Pr(PY_i = 57|\theta = 1.0)$ = dbinom(57,100,1.0) = 0

plot these probabilities as a function of theta

```
theta=seq(0,1,by=0.1)
p=c(0,4.107157e-31,3.738459e-16,1.306895e-
08,0.0002285792,0.03006864,0.06672895,0.001853172,1.003535e-07,9.395858e-18,0)
plot(theta,p,type='h',ylab="Prob(y=57|theta,n=100)")
```



c) Now suppose you originally had no prior information to believe one of these θ-values over another, and so $Pr(\theta = 0.0) = Pr(\theta = 0.1) = \cdots = Pr(\theta = 0.9) = Pr(\theta = 1.0)$. Use Bayes' rule to compute $p(\theta | \sum_{i=1}^{100} Y\, i = 57)$ for each θ-value. Make a plot of this posterior distribution as afunction of θ.

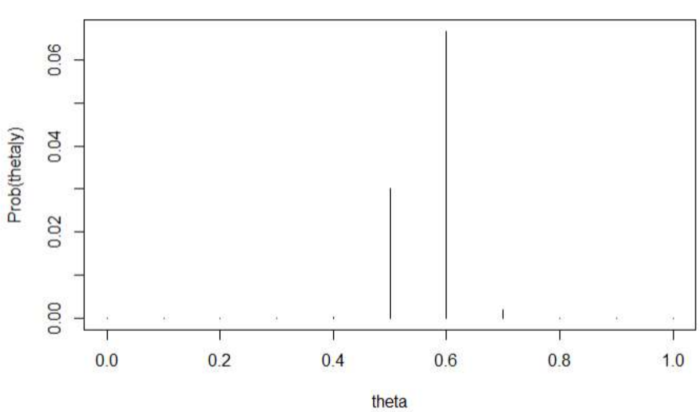
solution: $Pr(\theta = 0.0) = Pr(\theta = 0.1) = \cdots = Pr(\theta = 0.9) = Pr(\theta = 1.0)$=1/11

pr$(\theta = 0 | \sum_{i=1}^{100} Y\, i = 57)$= pr$(\theta = 0)$*$Pr(PY_i = 57 | \theta = 0.0)$/pr(Y)

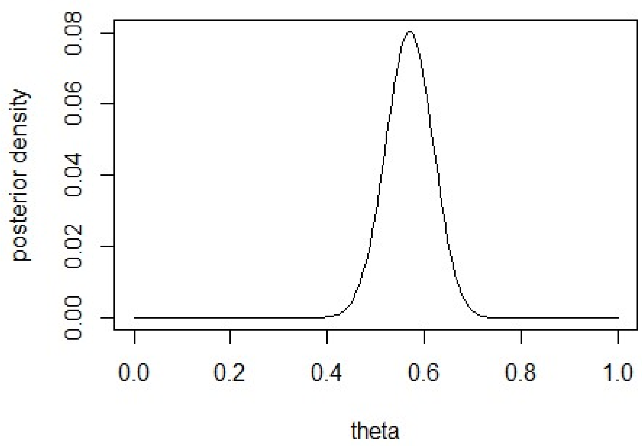
pr$(\theta \big| \sum_{i=1}^{100} Y\, i = 57)$ is proportional to Pr(θ)*Pr$(PY_i = 57 | \theta)$

plotted the posterior distribution as function of theta

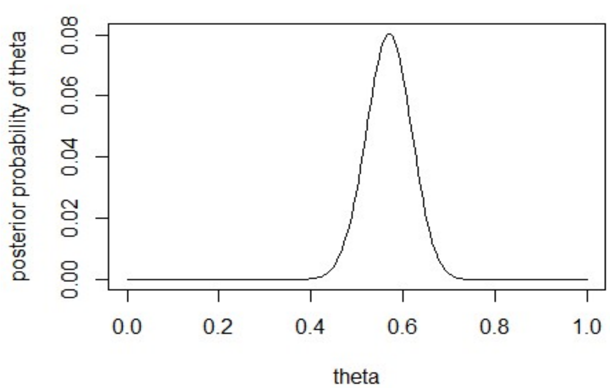
```
theta=seq(0,1,by=0.1)
p= c(0,4.107157e-31,3.738459e-16,1.306895e-
08,0.0002285792,0.03006864,0.06672895,0.001853172,1.003535e-07,9.395858e-18,0)
plot(theta,p,type='h',ylab="Prob(theta|y)")
```



```r
theta=seq(0,1,length=200)
prior_prob=1
post_prob_unwieghted=(choose(100,57))*(theta)^(57)*(1-theta)^(43)*1
plot(theta,post_prob_unwieghted,type='l',ylab="posterior density")
```




```r
theta=seq(0,1,length=200)
post_prob=dbeta(theta,58,44)
plot(theta,post_prob_unwieghted,type='l',ylab="posterior probability of theta")
```



Discussion of relationships among the plots.

we can see that a Beta exact posterior, we get more accuracy in the posterior estimate of theta, although this more precise estimate is pretty close to what we got with discrete priors. With discrete priors and the uniform specification, we got theta=0.6 as the posterior estimate for theta, but with exact beta posterior, we got slightly smaller than 0.6.

3.3

Given that $\theta_A\ gamma(120{,}10) and \theta_B\ gamma(12{,}1)$

we know from conjugacy that if the prior belongs to gamma family, then posterior will also belong to gamma family if the sampling model is poisson model.

The posterior distributions are: for θ_A: $\theta_A|y_A = gamma(120 + \sum_{i=1}^{10} y_A\, i = 237{,}10 + 10 = 20)$

for θ_B: $\theta_B|y_B = gamma(12 + \sum_{i=1}^{13} y_A\, i = 125{,}1 + 13 = 14)$

sy1=117 sy2=113

a_A=120 ;b_A=10 - # prior parameters for Theta_A

a_B=12 ;b_B=1 # prior parameters for Theta_B

n1=10 ; n2=13 # data for typeA # data for type B

Posterior_Mean_A = ((120+117) / (10+10)) =11.85 # Posterior mean for theta A

Posterior_Variance_A=((120+117)/(20)^2)=0.5925 # Variance for theta A

CI_A=qgamma(c(0.025,0.975) ,(a_A+sy1),(b_A+n1)) =10.38924 13.40545

posterior 95% CI for theta A

Posterior_Mean_B = ((a_B+sy2) /(b_B+n2))=8.92857 # Posterior mean for theta B

Posterior_Variance_B=((a_B+sy2 −1)/(b_B+n2)^2) =0.6326531 # Variance for theta B

CI_B=qgamma(c (0.025, 0.975) , (a_B+sy2) , (b_B+n2))= 7.432064 10.560308 # posterior 95% CI for theta B

b)

R-CODE

```
z <- seq(1, 50)

x <- 12*z

y <- z

a_post_new_B <- x + sum(y_b)

b_post_new_B<- y + length(y_b)

expectation.posterior_B <- a_post_new_B/b_post_new_B

expectation.posterior_B
```

```
[1]  8.928571  9.133333  9.312500  9.470588  9.611111  9.736842  9.850000  9.952381 10.045455
10.130435
[11] 10.208333 10.280000 10.346154 10.407407 10.464286 10.517241 10.566667 10.612903 10.656250
10.696970
[21] 10.735294 10.771429 10.805556 10.837838 10.868421 10.897436 10.925000 10.951220 10.976190
11.000000
[31] 11.022727 11.044444 11.065217 11.085106 11.104167 11.122449 11.140000 11.156863 11.173077
11.188679
[41] 11.203704 11.218182 11.232143 11.245614 11.258621 11.271186 11.283333 11.295082 11.306452
11.317460
```

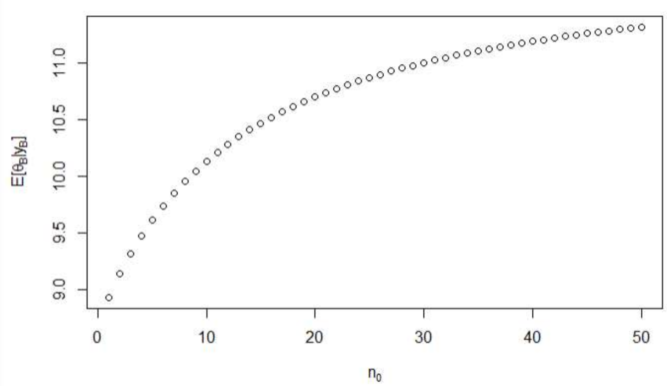
plot(z,

expectation.posterior_B,

xlab = expression(n[0]),

ylab = expression(paste("E[", theta[B], "|", y[B], "]", sep = "")))

abline(h = 11.85, lty = 2)



Given that we believe that the prior mean of θB is 12, we would need to have very strong prior beliefs about θB, corresponding to n0 ≈ 50, for the posterior expectation of θB to be close to that of θA.

C)

Based on the information that type B mice are related to type A mice, knowledge about population A should tell us something about population B. If knowledge about type A mice should tell us something about type B mice, then it does not make sense to have p (θA, θB) = P (θA) P(ΘB)because it indicates that our prior beliefs about the tumor rate among type A mice are independent of our prior beliefs about the tumor rate among type B mice. If we know the two types of mice are related, our prior beliefs about the tumor rates in type A and B mice should not be independent.

3.4)

a)

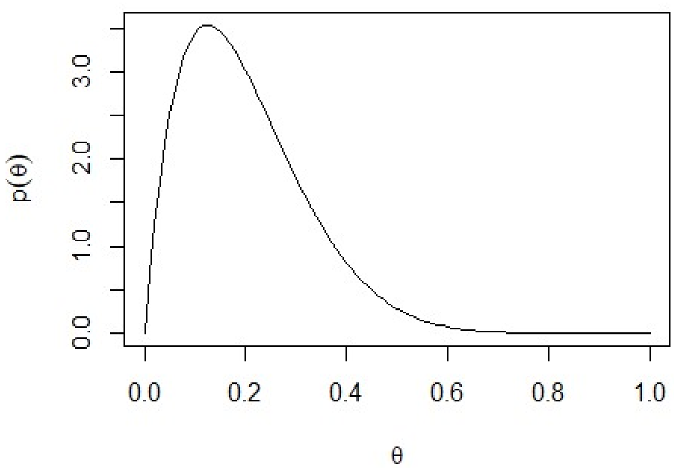
Assuming the teens act independently and have same chance of recidivism. The sampling model here is Binomial distribution. $Y = Binomial(n, \theta)$ = $Y = Binomial(43, \theta)$

a)using a beta(2,8) prior for θ: We know from conjugacy that if prior belongs to beta distribution, then posterior will also belong to beta distribution if the sampling model is binomial model.

$\theta|y \sim$ beta(a+y,b+n-y)= beta(2+15,8+43-15)=beta(17,36)

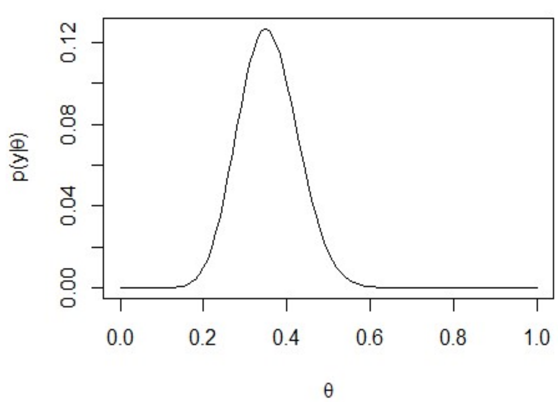
plots

```
y=15;n=43;a1=2;b1=8
theta=seq(0,1,by=0.01)
plot(theta,dbeta(theta,a1,b1),type='l',xlab=expression(theta),ylab=expression(p(theta
)))
```



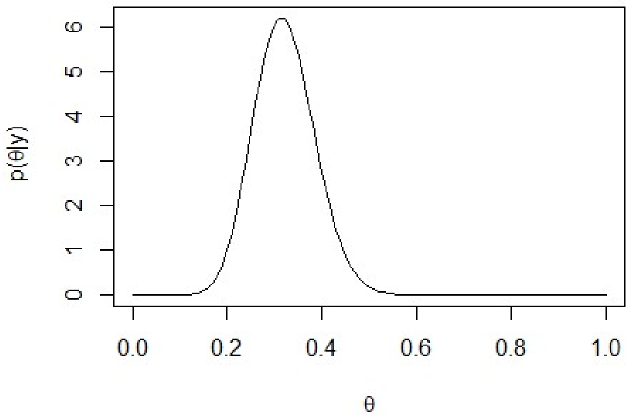
Sampling model(binomial Distribution)

```
plot(theta,dbinom(y,n,theta),type='l',xlab=expression(theta),ylab=expression(paste('p
(y|',theta,')')))
```



Plot of posterior model-Beta distribution

```
plot(theta,dbeta(theta,17,36),type='l',xlab=expression(theta),ylab=expression(paste('
p(',theta,'|y)')))
```



#calculation of posterior means, modes, standard deviation and 95% CI for beta prior(2,8)

```
a=a1;b=b1;y=15;n=43
posterior_mean=(a+y)/(a+y+b+n-y)
posterior_mean

## [1] 0.3207547

posterior_mode=(a+y-1)/(a+y+b+n-y-1)
posterior_mode

## [1] 0.3076923

posterior_standrad_dev=sqrt((a+y)*(b+n-y)/((a+y+b+n-y)^2*(a+y+b+n-y-1)))
posterior_standrad_dev

## [1] 0.06472889

CI_Interval=qbeta(c(0.025,0.975),a+y,b+n-y)
CI_Interval

## [1] 0.2032978 0.4510240
```

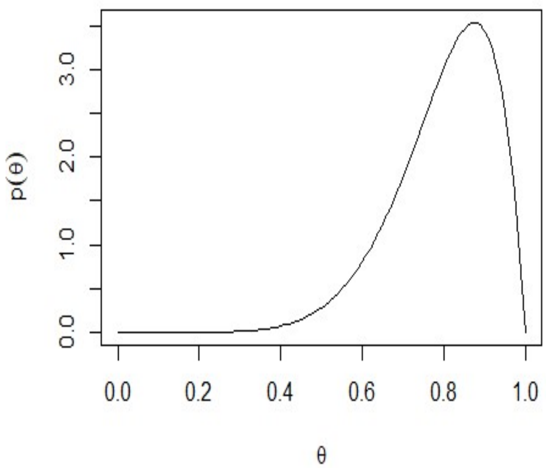
b) using a beta(8,2) prior for θ: We know from conjugacy that if prior belongs to beta distribution, then posterior will also belong to beta distribution if the sampling model is binomial model.

$\theta|y \sim$ beta(a+y,b+n-y)= beta(8+15,2+43-15)=beta(23,30)

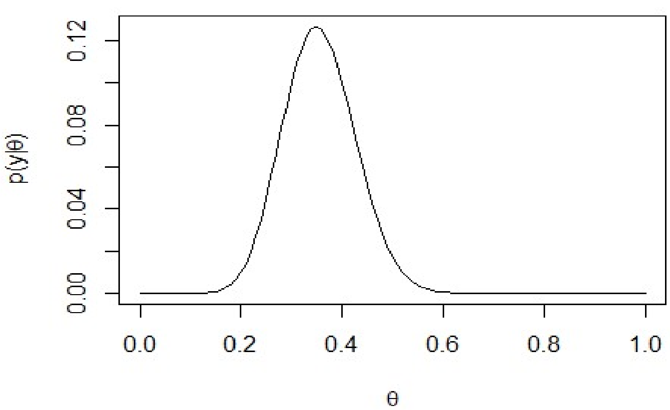
plots

```
y=15;n=43;a2=8;b2=2
theta=seq(0,1,by=0.01)
```

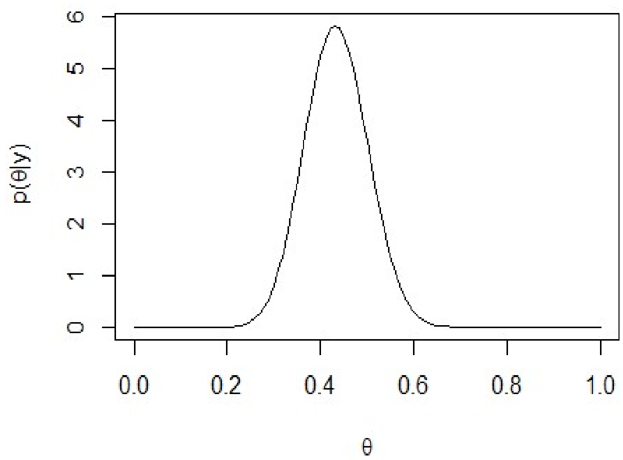
```
plot of prior distribution

plot(theta,dbeta(theta,a2,b2),type='l',xlab=expression(theta),ylab=expression(p(theta
)))
```




```
Plot of sampling model

plot(theta,dbinom(y,n,theta),type='l',xlab=expression(theta),ylab=expression(paste('p
(y|',theta,')')))
```




```
Plot of posterior distribution

plot(theta,dbeta(theta,23,30),type='l',xlab=expression(theta),ylab=expression(paste('
p(',theta,'|y)')))
```



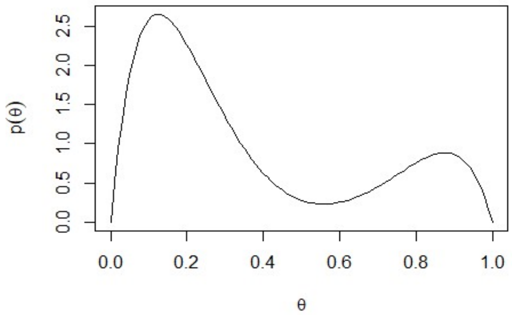
calculation of posterior means, modes, standard deviation and 95% CI for beta prior(8,2)

```
y=15;n=43;a=8;b=2
posterior_mean=(a+y)/(a+y+b+n-y)
posterior_mean

## [1] 0.4339623

posterior_mode=(a+y-1)/(a+y+b+n-y-1)
posterior_mode

## [1] 0.4230769

posterior_standrad_dev=sqrt((a+y)*(b+n-y)/((a+y+b+n-y)^2*(a+y+b+n-y-1)))
posterior_standrad_dev

## [1] 0.0687301

CI_Interval=qbeta(c(0.025,0.975),a+y,b+n-y)
CI_Interval

## [1] 0.3046956 0.5679528
```

There is a 95% chance that the true θ(proportion of teen recidivism) is between 30% and 57% when beta(8,2)prior

plot the prior distribution of mixture 75-25% mixture of beta(2,8) and beta(8,2)

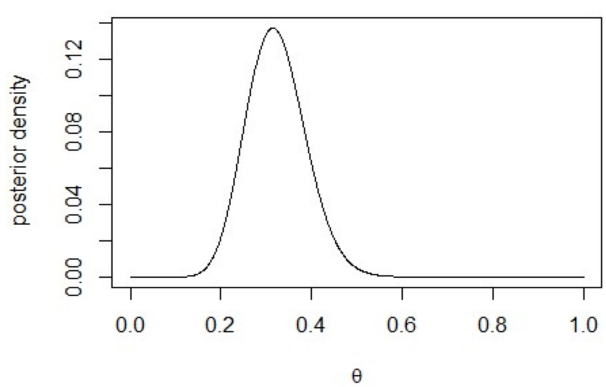
```
a1=2;b1=8
y=15;n=43;a2=8;b2=2
theta=seq(0,1,by=0.01)

plot(theta,0.75*dbeta(theta,a1,b1)+0.25*dbeta(theta,a2,b2),type='l',xlab=expression(t
heta),ylab=expression(p(theta)))
```



since 75% weightage is given to beta prior(2,8), hence the mixture prior distribution is more inclinded towrads the same prior belief, as we can see a peak on the region between θ 0.1 to 0.3 which looks similar to prior plot of beta prior (2,8).

3.4 d)iii)

```
p1=function(theta){
  return(18*choose(43,15)*(3*theta^16*(1-theta)^35+theta^22*(1-theta)^29))

}

theta=seq(0,1,by=0.001)

plot(theta,p1(theta),type='l',xlab=expression(theta),ylab=expression('posterior
density'))
```

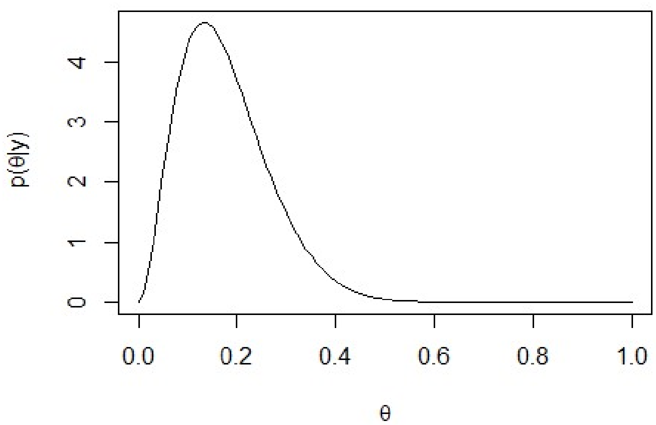



```
mode=theta[which.max(p1(theta))]
mode

## [1] 0.314
```

the mode found here lies between posterior modes for beta prior(2,8) and beta prior(8,2) but closer to posterior mode of beta prior(2,8).

3.7 a)

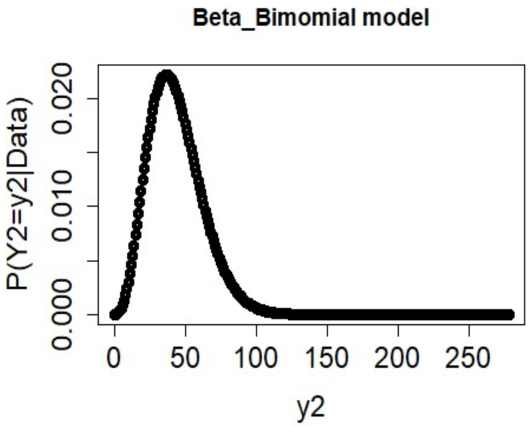
```
theta=seq(0,1,by=0.01)
plot(theta,dbeta(theta,3,14),type='l',xlab=expression(theta),ylab=expression(paste('p
(',theta,'|y)')))
```




```
# Beta Binomial Predictive distribution function

##reference was taken from : https://rpubs.com/FJRubio/BetaBinomialPred
```

3.7)c

```
# Beta Binomial Predictive distribution function
BetaBinom <- Vectorize(function(rp){
log.val <- lchoose(np, rp) + lbeta(rp+a+r,b+n-r+np-rp) - lbeta(a+r,b+n-r)
return(exp(log.val))
})

n <- 15; r <- 2; a <- 3; b <- 14; np <- 278
plot(0:278,BetaBinom(0:278),type="b",xlab="y2",ylab="P(Y2=y2|Data)", main = "Beta_Bimomial
model",cex.axis= 1.5,cex.lab=1.5,lwd=4)
```



Obtain the mean and standard deviation of Y2, given Y1 = 2.

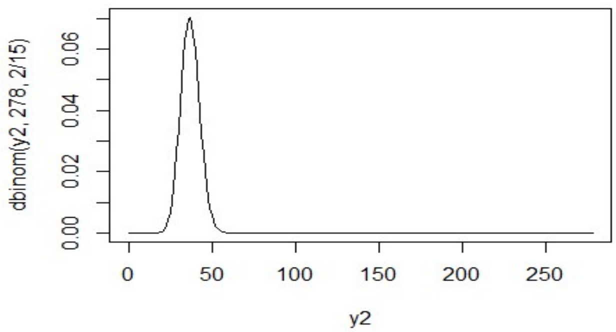
```
# Beta Binomial Predictive distribution function
a=3;b=14;n=278
mean=n*a/(a+b)
variance=n*a*b*(a+b+n)/(((a+b)^2)*(a+b+1))

std_dev=sqrt(variance)
std_dev

## [1] 25.73196

Mean

## [1] 49.05882
```

3.7(d)

Calculation of mean from classical model: n*theta'=278*2/15=37.067

Variance=n*theta'(1-theta')=32.1244

```
y2=seq(0,278,1)

plot(y2,dbinom(y2,278,2/15),type='l')
```



Since, we have less data, hence Bayesian estimation which depends upon the prior belief would provide better result, as in less data classical method would mislead us.

Here, we can say that the estimation from betabinom predictive model, we have more accurate prediction. From the results we can see that, the mean value for y2 given by betabinom predictive model is 49.05882 which seems reasonable from the plots and mean from the classical method is 37.