Bayesian Computing

Name :- Preksha Ashok Patel
Sapid :- 60004210126
Branch :- Computer Engineering
Div :- C2-1.

Experiment no. 2

Aim :- To implement a Discrete Prior, Beta Prior, Histogram Prior, Prediction model using R

Theory :

Discrete Prior :-

A discrete Prior is a probability distribution over a finite or countably infinite set of possible values. It represents the belief about a parameter before any data ~~if~~ is observed. Discrete Priors are commonly used for categorical variables, such as gender or political affiliation.

Mathematically, a discrete prior for a ~~person~~ parameter can be represented as follows : $P(\theta = \theta_i) = p_i$

where θ_i is one of the possible values of θ and p_i is the probability that $\theta = \theta_i$

Beta Prior :-

A beta prior is a probability distribution over the interval $[0,1]$. It is commonly used as a prior for parameters that represent proportions or probabilities. The beta prior is characterized by 2 hyperparameters, α and β, which control the shape of the distribution. Mathematically, the beta prior is defined as follows :-

$$P(\theta \mid \alpha, \beta) = B(\alpha, \beta)\, \theta^{\alpha-1} (1-\theta)^{\beta-1}$$

where $B(\alpha, \beta)$ is the beta function.

The beta prior is a conjugate prior for the binomial distribution.

Page No.
Date / /

Histogram Prior :-
A Histogram Prior is a probability distribution that takes the form of a histogram. It can be used as a prior for any type of variable. but it is ~~of ten used~~ often used for continuous variable. The histogram prior is specified by a sequence of bin boundaries and associated probabilities. Mathematically, a histogram prior is defined as :-

$P(x \in b_i) = p_i$

where b_i is i^{th} bin and p_i is the probability that x falls in bin b_i.

Prediction Model :-
A prediction model is a statistical model that is used to predict future observations or events. In Bayesian computing, prediction models are typically based on posterior distributions. The posterior distribution provides a probability distribution for the unknown parameters of the model, which can then be used to make predictions about future observations.

Conclusion :-
In this experiment, we learned that how to implement different types of priors and posterior distributions for a binomial parameter p using R. we also learned how to construct and simulate ~~for~~ from the predictive distribution for a future observation. we compared the results of using a discrete, beta and histogram prior for p and how they affected the posterior and predictive distributions. we observed how prior information & data influenced shape of and location of posterior distribution and how the posterior distribution reflected the uncertainty about p. we also observed how predict distribution incorporated both the uncertainty about p & variability of future observation.

SAP ID.: 60004210126 Name: Preksha Ashok Patel

Batch: C2-1

Subject: Bayesian Computing Laboratory Semester: VII

Experiment No. 2

Aim:

To implement a Discrete Prior, Beta Prior, Histogram Prior, Prediction Model using R.

Code:

Loading the dataset

```
library('LearnBayes')
```

Using a Discrete Prior

```
library('LearnBayes') p <- seq(0.05, 0.95, by =
0.1) prior <- c(1, 5.2, 8, 7.2, 4.6, 2.1, 0.7, 0.1, 0,
0) prior <- prior / sum(prior) plot(p, prior, type =
"h", ylab="Prior Probability")
```

The posterior for p

```
data <- c(11, 16) post <-

pdisc(p, prior, data)

round(cbind(p, prior,

post),2)


library(lattice)

PRIOR <- data.frame("prior", p, prior)

POST <- data.frame("posterior", p, post)
```

```
names(PRIOR) <- c("Type", "P", "Probability") names(POST) <-
c("Type","P","Probability") data <- rbind(PRIOR, POST) xyplot(Probability ~ P |
Type, data=data, layout=c(1,2), type="h", lwd=3, col="black")
```

Using a Beta Prior

```
quantile2 <- list(p=.9, x=.5)

quantile1 <- list(p=.5, x=.3)

(ab <- beta.select(quantile1,quantile2))
```

Bayesian triplot

```
a <- ab[1] b <- ab[2] s <- 11 f <- 16 curve(dbeta(x, a + s, b + f), from=0, to=1,
xlab="p", ylab="Density", lty=1, lwd=4) curve(dbeta(x, s + 1, f + 1),
add=TRUE, lty=2, lwd=4) curve(dbeta(x, a, b), add=TRUE, lty=3, lwd=4)
legend(.7, 4, c("Prior", "Likelihood", "Posterior"), lty=c(3, 2, 1), lwd=c(3, 3, 3))
```

Posterior summaries

```
1 - pbeta(0.5, a + s, b + f)

qbeta(c(0.05, 0.95), a + s, b + f)
```

Simulating from posterior

```
ps <- rbeta(1000, a + s, b + f)
hist(ps, xlab="p")
```

```
sum(ps >= 0.5) / 1000

quantile(ps, c(0.05, 0.95))
```

Using a Histogram Prior

```
midpt <- seq(0.05, 0.95, by = 0.1) prior <- c(1, 5.2, 8, 7.2, 4.6, 2.1, 0.7, 0.1, 0, 0)

prior <- prior / sum(prior) curve(histprior(x, midpt, prior), from=0, to=1,

ylab="Prior density", ylim=c(0, .3))

curve(histprior(x,midpt,prior) * dbeta(x, s + 1, f + 1), from=0, to=1, ylab="Posterior
density")

p <- seq(0, 1, length=500) post <- histprior(p, midpt,
prior) * dbeta(p, s + 1, f + 1) post <- post / sum(post)
ps <- sample(p, replace = TRUE, prob = post) hist(ps,
xlab="p", main="")
```

Prediction Model

Discrete prior approach

```
p <- seq(0.05, 0.95, by=.1) prior <- c(1, 5.2,
8, 7.2, 4.6, 2.1, 0.7, 0.1, 0, 0) prior <- prior /
sum(prior) m <- 20 ys <- 0:20

pred <- pdiscp(p, prior, m, ys)
cbind(0:20, pred)
```

Continuous prior approach

```
ab <- c(3.26, 7.19) m <-
20 ys <- 0:20 pred <-
pbetap(ab, m, ys)
```

Simulating predictive distribution

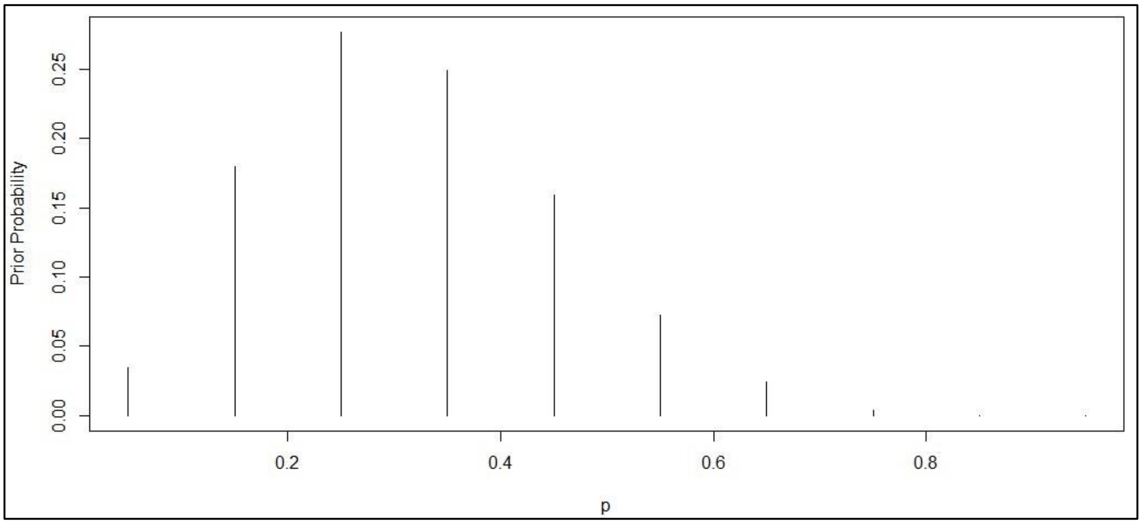
```
p <- rbeta(1000, 3.26, 7.19)

y <- rbinom(1000, 20, p)

table(y)


freq <- table(y) ys <- as.integer(names(freq)) predprob <- freq /

sum(freq) plot(ys, predprob, type="h", xlab="y", ylab="Predictive

Probability")


dist <- cbind(ys, predprob)
```

Construction of a prediction interval

```
covprob     <-     .9
discint(dist, covprob)
```

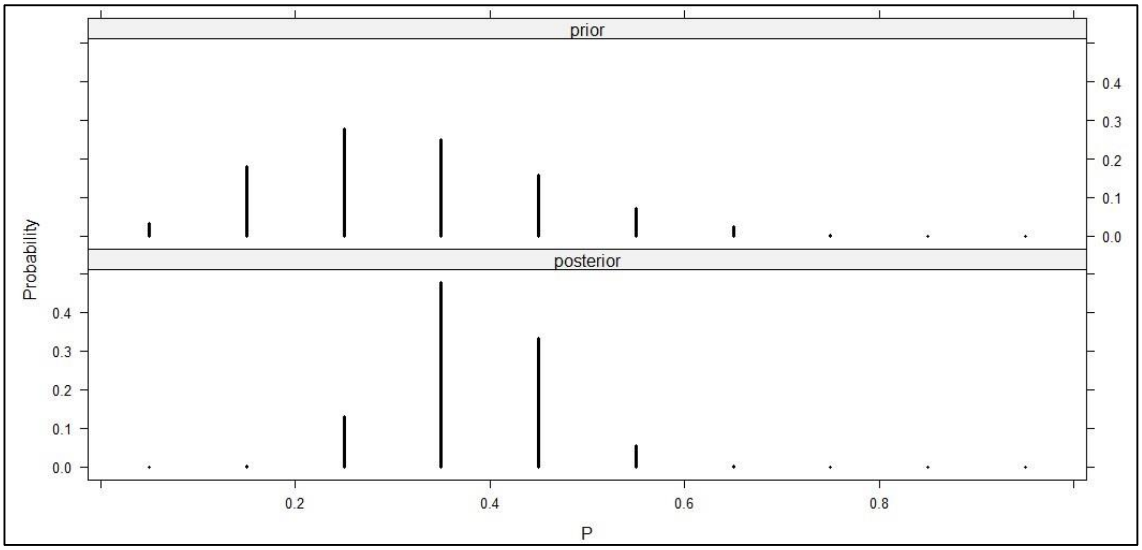
Output:

Using a Discrete Prior



The posterior for p

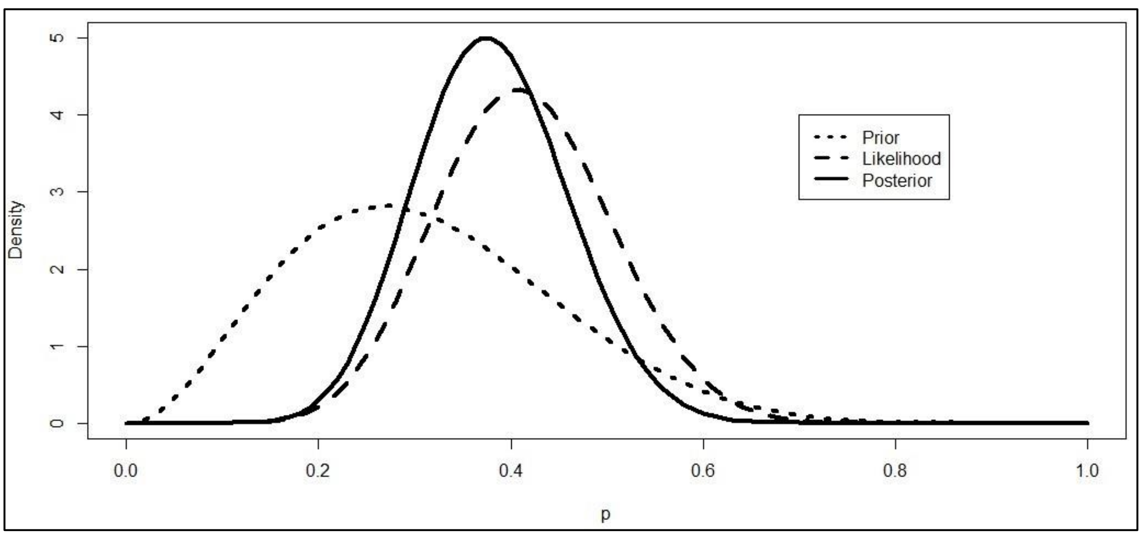
	p	prior	post
[1,]	0.05	0.03	0.00
[2,]	0.15	0.18	0.00
[3,]	0.25	0.28	0.13
[4,]	0.35	0.25	0.48
[5,]	0.45	0.16	0.33
[6,]	0.55	0.07	0.06
[7,]	0.65	0.02	0.00
[8,]	0.75	0.00	0.00
[9,]	0.85	0.00	0.00
[10,]	0.95	0.00	0.00



Using a Beta Prior

```
3.26 7.19
```

Bayesian triplot

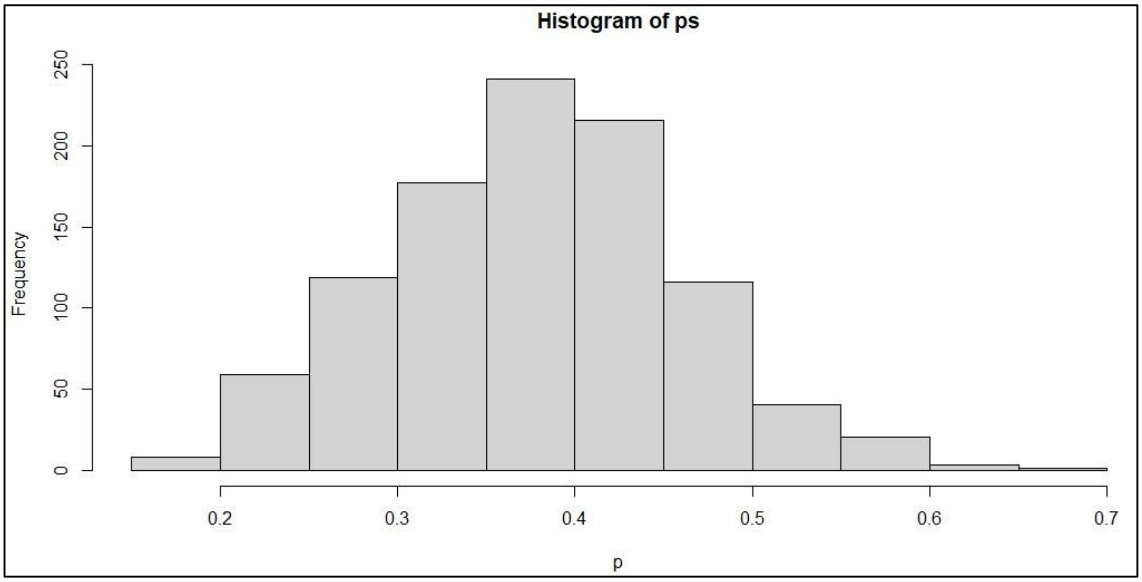


Posterior summaries

```
0.0690226
```

```
0.2555267 0.5133608
```

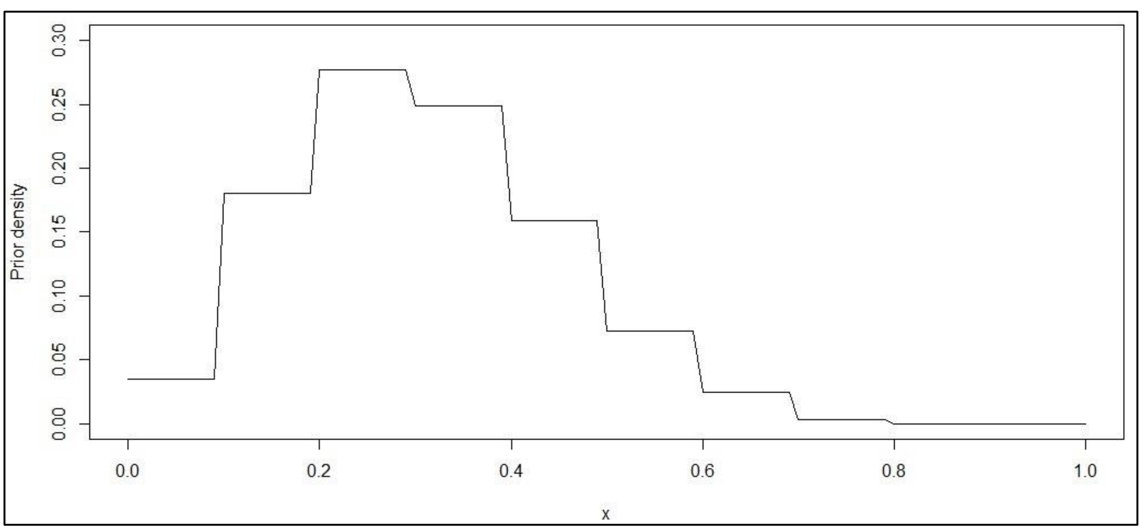
Simulating from posterior

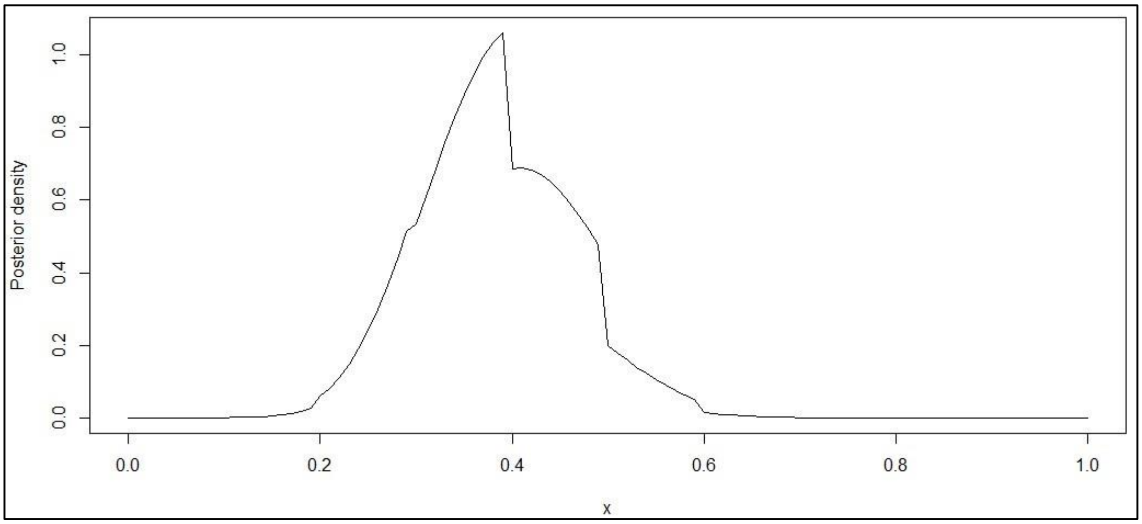


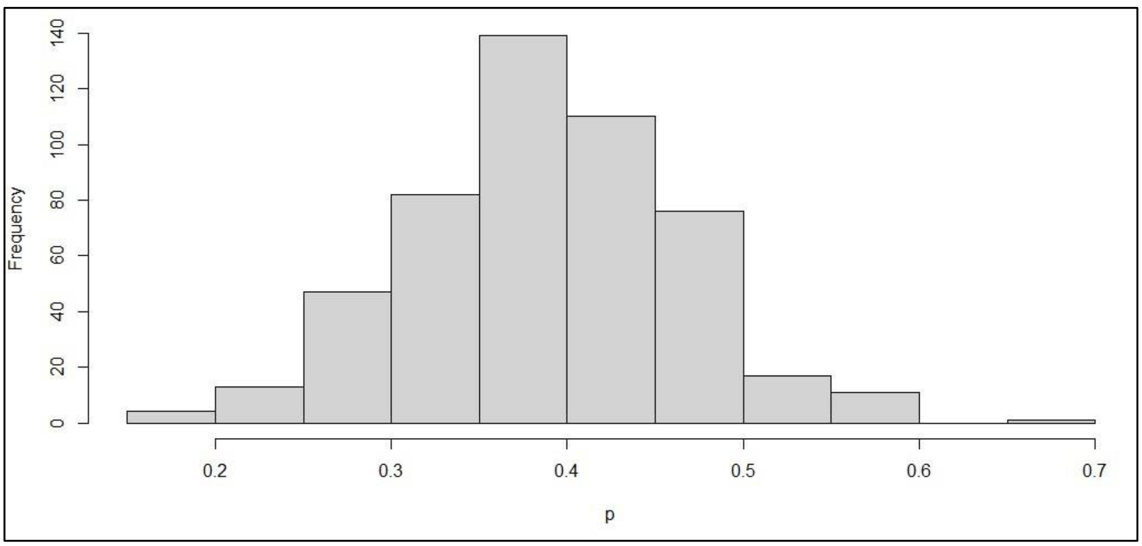

```
0.064
```

```
        5%       95%
0.2373968 0.5094293
```

Using a Histogram Prior




Posterior density
x


Frequency
p

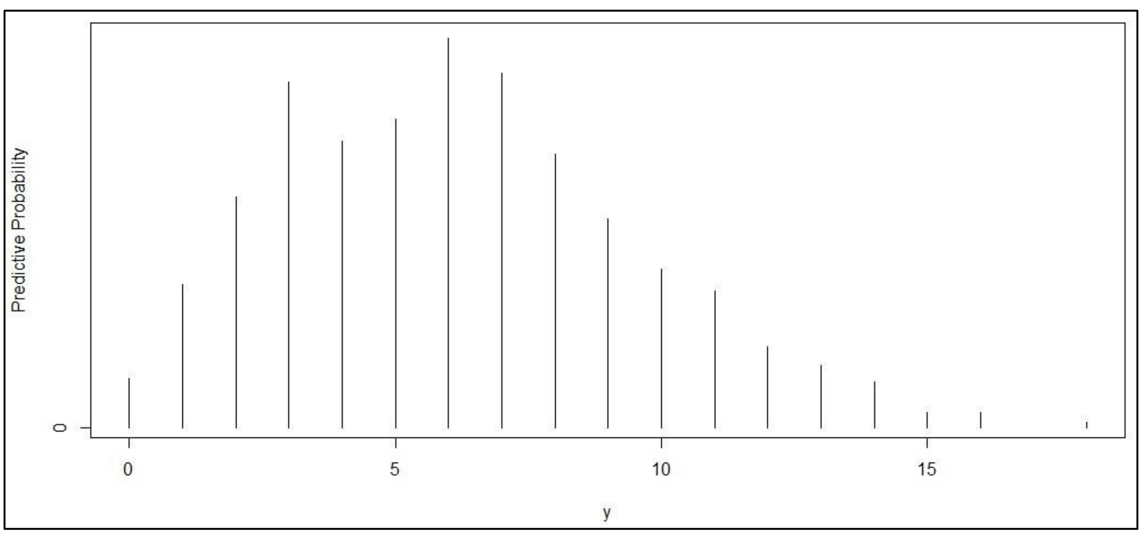
Prediction Model

Discrete prior approach

```
                pred
 [1,]  0 2.030242e-02
 [2,]  1 4.402694e-02
 [3,]  2 6.894572e-02
 [4,]  3 9.151046e-02
 [5,]  4 1.064393e-01
 [6,]  5 1.124487e-01
 [7,]  6 1.104993e-01
 [8,]  7 1.021397e-01
 [9,]  8 8.932837e-02
[10,]  9 7.416372e-02
[11,] 10 5.851740e-02
[12,] 11 4.383668e-02
[13,] 12 3.107700e-02
[14,] 13 2.071698e-02
[15,] 14 1.284467e-02
[16,] 15 7.277453e-03
[17,] 16 3.667160e-03
[18,] 17 1.575535e-03
[19,] 18 5.381536e-04
[20,] 19 1.285179e-04
[21,] 20 1.584793e-05
```

Simulating predictive distribution

```
y
  0   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  18
 16  46  74 111  92  99 125 114  88  67  51  44  26  20  15   5   5   2
```



Construction of a prediction interval

```
$prob
   11
0.911

$set
 1  2  3  4  5  6  7  8  9 10 11
 1  2  3  4  5  6  7  8  9 10 11
```