Assignment – 2

Name:- Pritam Priyadarshi

Roll No.:- 230793

Part 1: A simple binomial model

The likelihood assumption :-

$$L(\theta|y) = \frac{10!}{y! \cdot (10-y)!} \theta^y \cdot (1-\theta)^{10-y}$$

The prior assumption :-

$$p(\theta) = \begin{cases} 1, & when\ 0 \leq \ \theta \leq 1 \\ 0, & when\ \theta < 0\ or\ \theta > 1 \end{cases}$$

The data: y = 7

The marginal likelihood: $\int L(\theta|y) \cdot p(\theta) \cdot d\theta = \frac{1}{11}$

1.1 $L(\theta|7) = \frac{10!}{7! \cdot (10-7)!} \theta^7 \cdot (1-\theta)^{10-7} = 120 \cdot \theta^7 \cdot (1-\theta)^3$

(a) $\theta = 0.75$

$$p(\theta|7) = \frac{(120 \cdot (0.75)^7 \cdot (0.25)^3) \cdot (1)}{{}^1/_{11}}$$

$$= \mathit{2.75310516}$$

(b) $\theta = 0.25$

$$p(\theta|7) = \frac{(120 \cdot (0.25)^7 \cdot (0.75)^3) \cdot (1)}{{}^1/_{11}}$$

$$= \mathit{0.03398895}$$

(c) $\theta = 1$

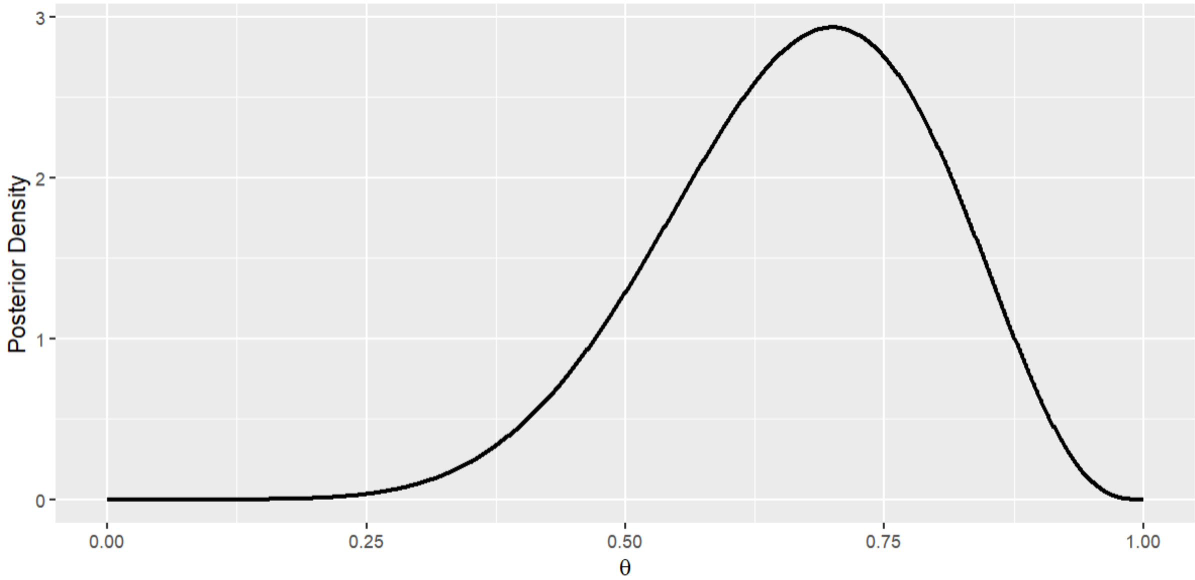
$$p(\theta|7) = \frac{(120 \cdot (1)^7 \cdot (0)^3) \cdot (1)}{{}^1/_{11}}$$

$$= \mathit{0}$$

1.2

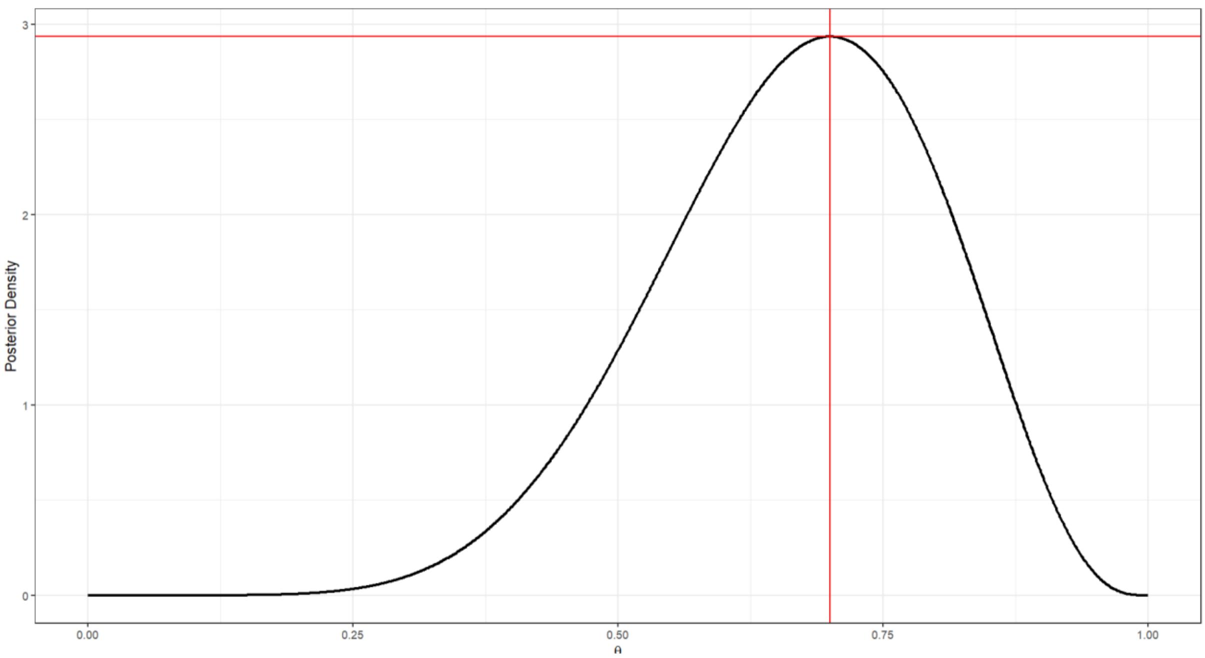
```
y <- 7
n <- 10
ml <- 1/11
theta <- seq(from=0, to=1, length=1000)
lkl <- dbinom(y, 10, theta)
likelihoods <- data.frame(theta=theta, lkl=lkl)
likelihoods$prior_density <- ifelse(theta<=1 & theta>=0, 1, 0)
likelihoods$posterior_density <- (likelihoods$lkl*likelihoods$prior_density)/
  ml
library(ggplot2)
ggplot(likelihoods, aes(x=theta, y=posterior_density))+
```

```r
  geom_line(size=1, color="black")+xlab(expression(theta))+
  ylab("Posterior Density")

```



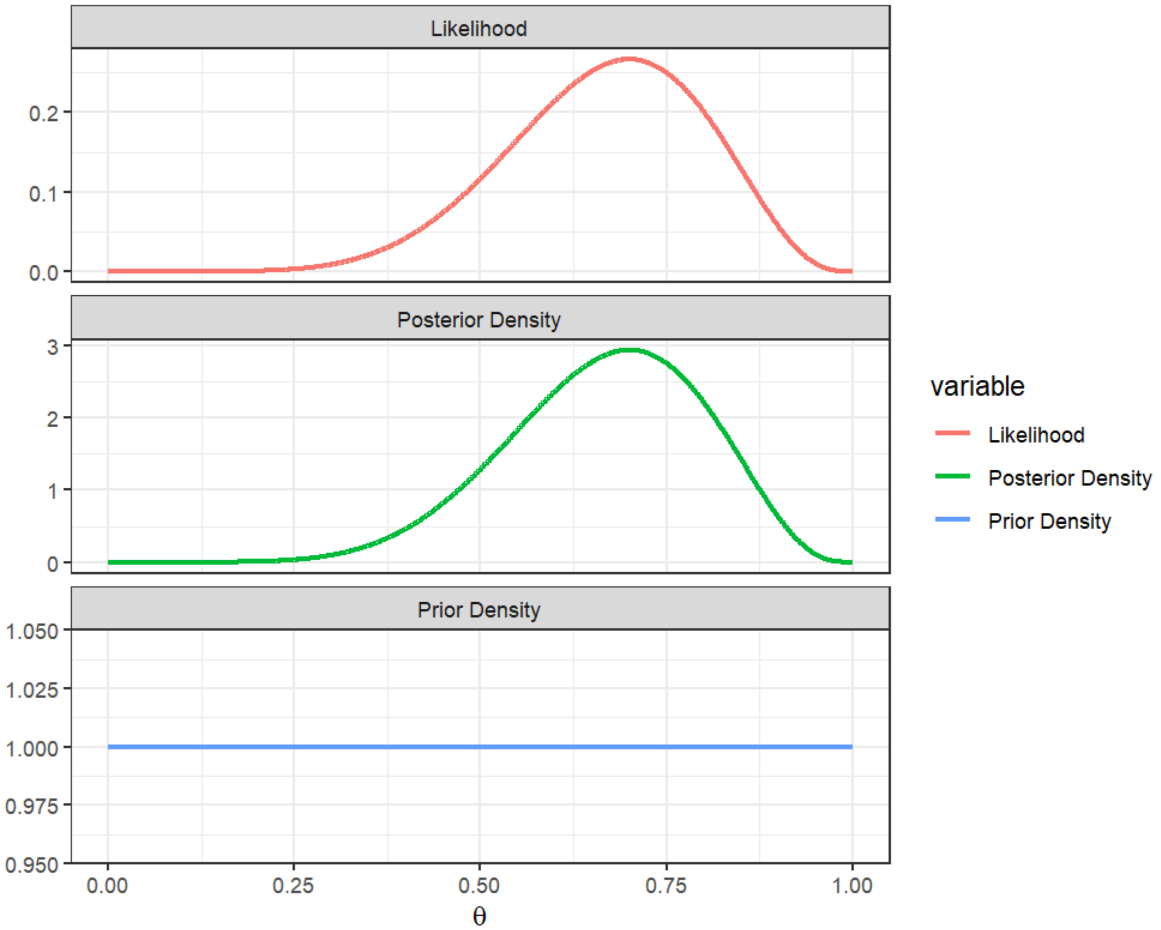
1.3

```r
post_max <- max(likelihoods$posterior_density)
theta_max <- likelihoods[likelihoods$posterior_density == post_max,c(theta)]

ggplot(likelihoods, aes(x=theta, y=posterior_density))+
  geom_line(size=1, color="black")+xlab(expression(theta))+
  ylab("Posterior Density")+
  geom_hline(yintercept = post_max, color="red")+
  geom_vline(xintercept = theta_max, color="red")+theme_bw()

```



1.4

```
library(reshape2)
df.lkl <- melt(likelihoods, id = c("theta"))
df.lkl$variable <- ifelse(df.lkl$variable == "lkl", "Likelihood",
                          ifelse(df.lkl$variable == "prior_density",
                                 "Prior Density", "Posterior Density"))
View(df.lkl)

ggplot(df.lkl, aes(x=theta, y=value, colour = variable))+
  geom_line(size=1)+xlab(expression(theta))+
  ylab("")+theme_bw()+facet_wrap(~variable, ncol=1, scales="free_y")

```



Part 2: A Gaussian model of reading

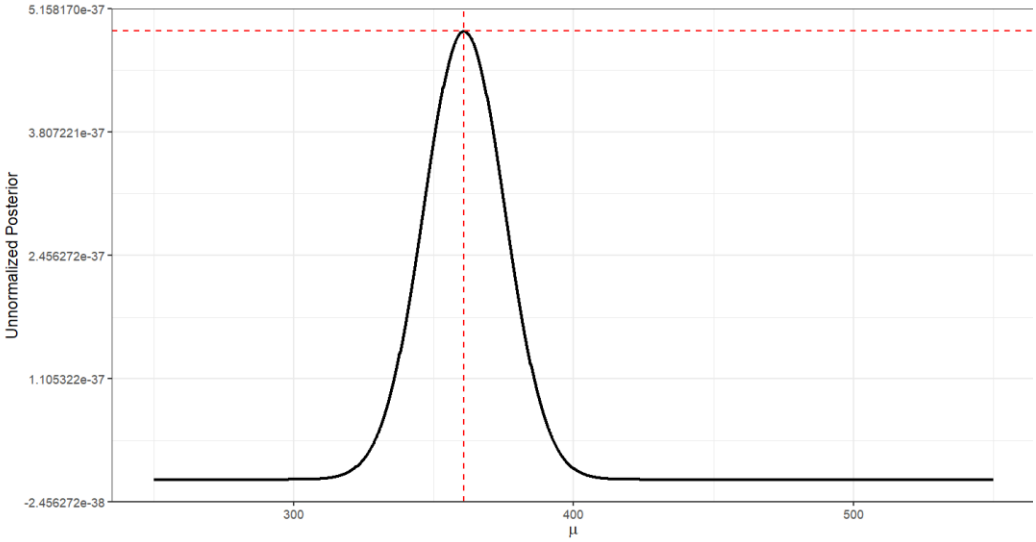
```r
# The likelihood assumption
y <- c(300, 270, 390, 450, 500, 290, 680, 450)
sd <- 50
mu <- seq(from=0, to=1000, by=0.25)

data <- data.frame(mu=mu, sigma=sd)
data$lkl <- rep(NA, length(mu))

for (i in 1:length(mu)) {
  data$lkl[i] <- prod(dnorm(y, mean = mu[i], sd = sd))
}

sigma <- 50
# mu ~ Normal(250,25)

# p(sigma) = 1 when sigma=50 and = 0 when sigma!=50
# p(mu) = dnorm(mu, 250, 25)

data$prior <- dnorm(mu, mean = 250, sd=25)
data$unnorm_posterior <- data$lkl*data$prior

```

2.1

```r
mu_given = c(300, 900, 50)

for (i in mu_given) {
  x <- data[data$mu == i, c("unnorm_posterior")]
  print(x)
}

# [1] 6.824248e-41
# [1] 0
# [1] 9.691374e-138

```

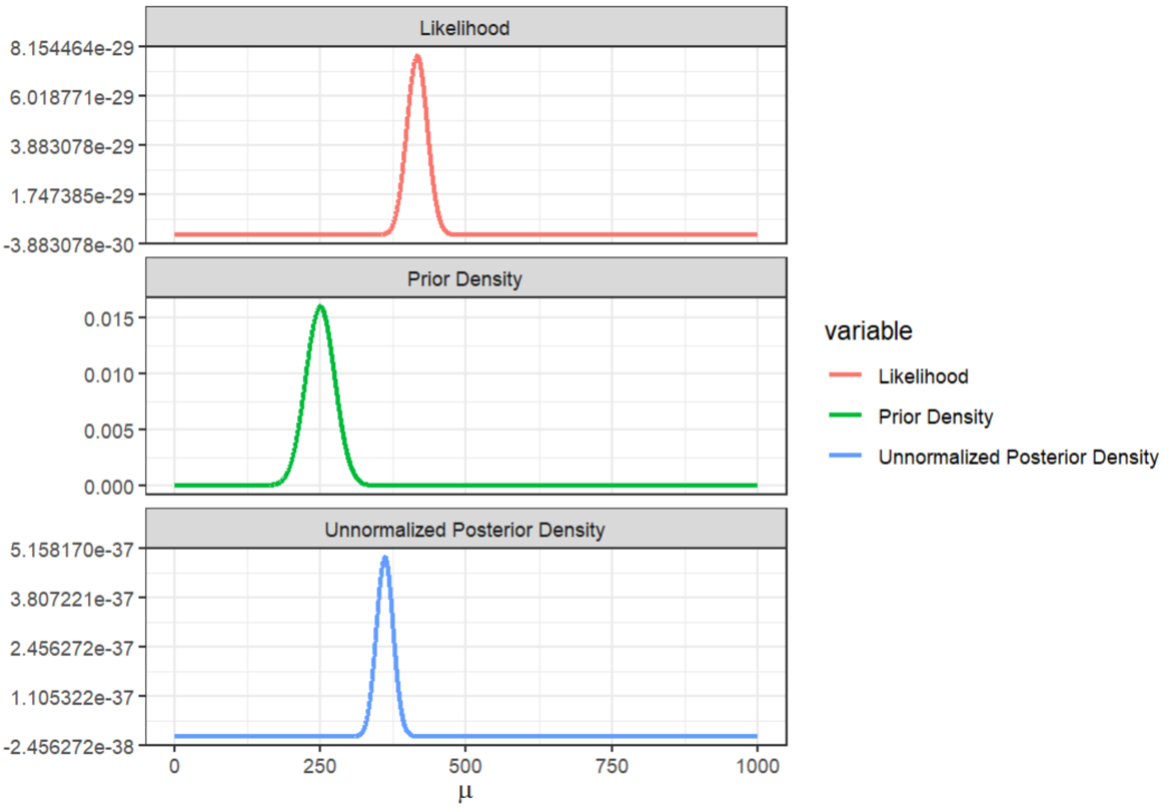
2.2

```r
post_max <- max(data$unnorm_posterior)
mu_max <- data[data$unnorm_posterior == post_max,c("mu")]

ggplot(data, aes(x=mu, y=unnorm_posterior))+
  geom_line(size=1, color="black")+
  scale_x_continuous(limits = c(250,550))+
  xlab(expression(mu))+ylab("Unnormalized Posterior")+
  geom_hline(yintercept = post_max, linetype="dashed", color="red")+
  geom_vline(xintercept = mu_max, linetype="dashed", color="red")+
  theme_bw()

```



2.3

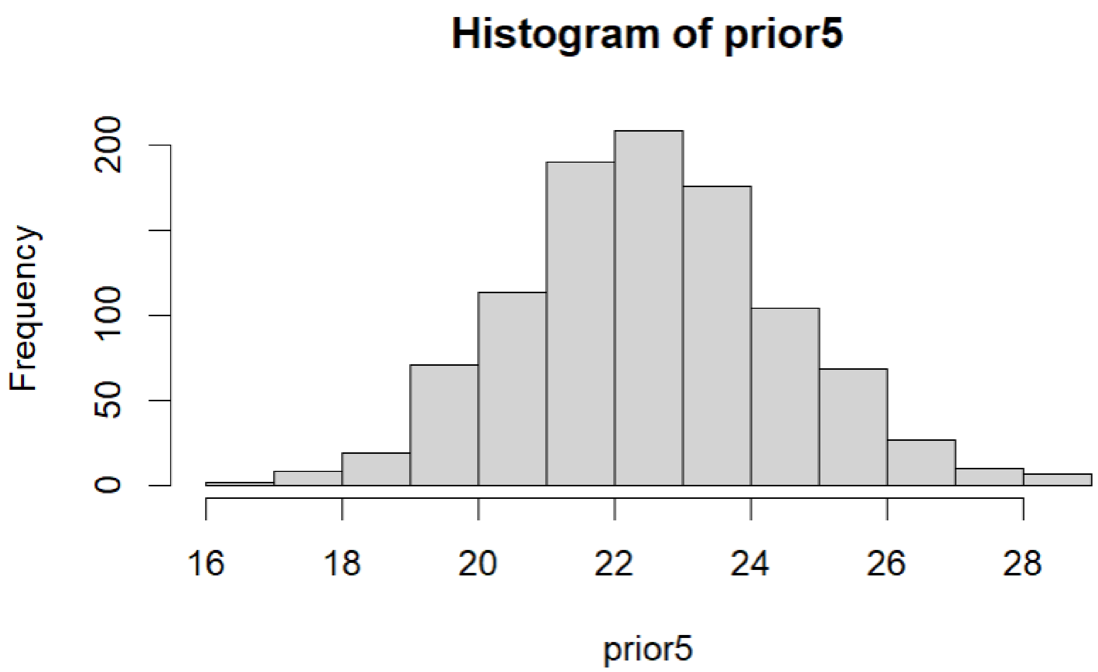
```
library(reshape2)
df.data <- melt(data, id = c("mu", "sigma"))
View(df.data)
df.data$variable <- ifelse(df.data$variable == "lkl", "Likelihood",
                           ifelse(df.data$variable == "prior",
                                  "Prior Density",
                                  "Unnormalized Posterior Density"))
View(df.data)

ggplot(df.data, aes(x=mu, y=value, colour = variable))+
  geom_line(size=1)+xlab(expression(mu))+
  ylab("")+theme_bw()+facet_wrap(~variable, ncol=1, scales="free_y")

```



Part 3: The Bayesian learning

3.1

```
# k = No. of accidents in a day
k <- c(25, 20, 23, 27)
# n = No. of days
n <- length(k)

# Prior on the parameter λ: λ ~ Gamma(40, 2)
# The posterior distribution of λ analytically:
# λ ~ Gamma(40+k, 3)

# The prior on λ to generate predictions for day 5
prior5 <- rgamma(1000, 40+sum(k), 2+n)
hist(prior5)

```



3.2

```
# Road accidents are predicted to happen on day 5
mean(prior5)
(40+sum(k))/(2+n)


# [1] 22.53353
# [1] 22.5

```

Part 4: Model building in the Bayesian framework

```
library(truncnorm)
dat <- read.table(
  "https://raw.githubusercontent.com/yadavhimanshu059/CGS698C/main/notes/Module-2/recognition.csv",
  sep=",",header = T)[,-1]
head(dat)

#      Tw       Tnw
# 1 285.0780 296.8060
# 2 267.5184 280.1157
# 3 289.9203 310.4417
# 4 399.0674 324.8276
# 5 359.9884 373.8152
# 6 403.3993 269.8220

sigma <- 60
mu<- seq(from=100, to=600, length=1000)

# NULL Hypothesis Model

delta_null <- 0

dat_null <- data.frame(mu=mu, sigma=sigma, delta_null=delta_null)

# likelihoods of words and non words
dat_null$lkl_w <- rep(NA, length(mu))
dat_null$lkl_nw <- rep(NA, length(mu))

for (i in 1:length(mu)) {
  dat_null$lkl_w[i] <- prod(dnorm(dat$Tw, mean=mu[i], sd=sigma))
  dat_null$lkl_nw[i] <- prod(dnorm(dat$Tnw, mean=mu[i]+delta_null, sd=sigma))
}

# now priors
dat_null$prior_mu <- dnorm(mu, 300, 50)

# since for null hypothesis, delta=0; thus prior or probability
# of this delta=1
dat_null$prior_delta <- 1

# posterior of Null Hypothesis
dat_null$post_unnorm <-
(dat_null$lkl_w*dat_null$lkl_nw*dat_null$prior_mu*dat_null$prior_delta)

View(dat_null)
library(ggplot2)
ggplot(dat_null, aes(x=mu, y=post_unnorm))+
  geom_line(size=1, color="black")+theme_bw()+
  xlab(expression(mu))+ylab("Unnormalized Posterior Null Hypothesis")

```