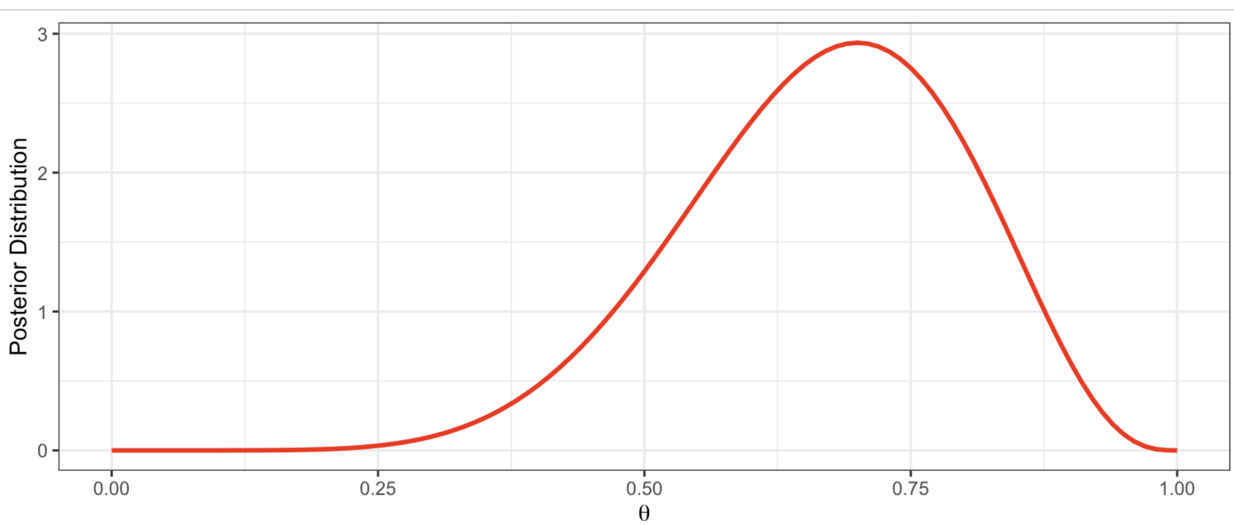
Name: Ishaan Jain
Roll Number: 230483
Part 1) 1.1)

```
posterior_0.75 <- 11*prod(dbinom(x=7, size=10, prob=0.75))
posterior_0.25 <- 11*prod(dbinom(x=7, size=10, prob=0.25))
posterior_1 <- 11*prod(dbinom(x=7, size=10, prob=1))
```

```
> cat("1.1) a) Posterior at 0.75:", posterior_0.75, "\n")
1.1) a) Posterior at 0.75: 2.753105
> cat("1.1) b) Posterior at 0.25:", posterior_0.25, "\n")
1.1) b) Posterior at 0.25: 0.03398895
> cat("1.1) c) Posterior at 1.00:", posterior_1, "\n")
1.1) c) Posterior at 1.00: 0
```

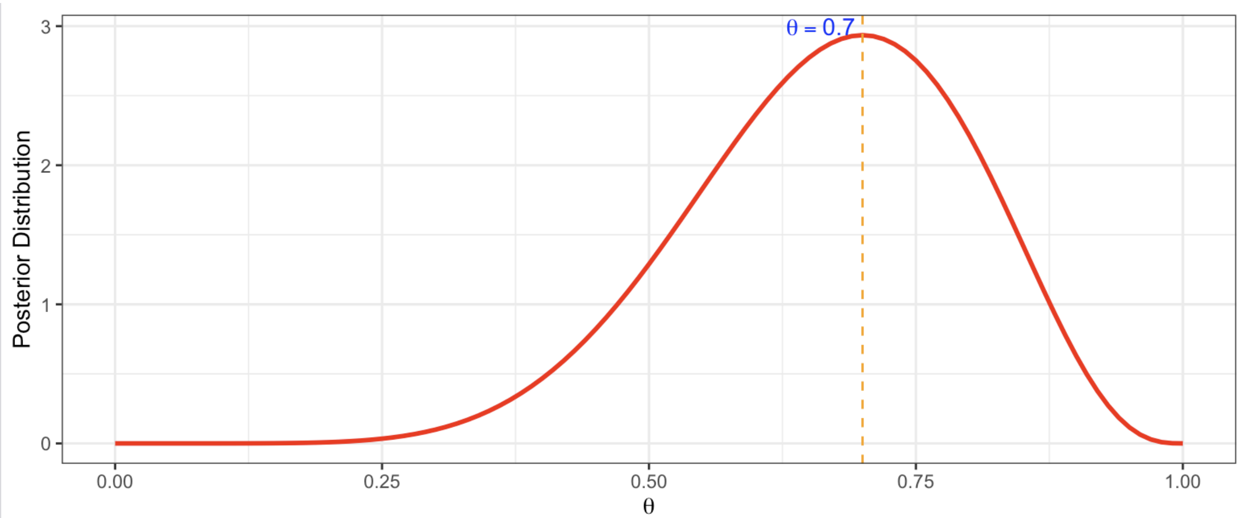
1.2)

```
theta <- seq(from=0, to=1, by=0.01)
posterior <- data.frame(theta=theta)
posterior$pst <- NA
for (i in 1:length(theta)){
  posterior$pst[i] <- 11*prod(dbinom(x=7, size=10, prob=theta[i]))
}
ggplot(posterior, aes(x=theta, y=pst))+geom_line(linewidth=1, color="red")+
  theme_bw() + xlab(expression(theta)) + ylab("Posterior Distribution")
```



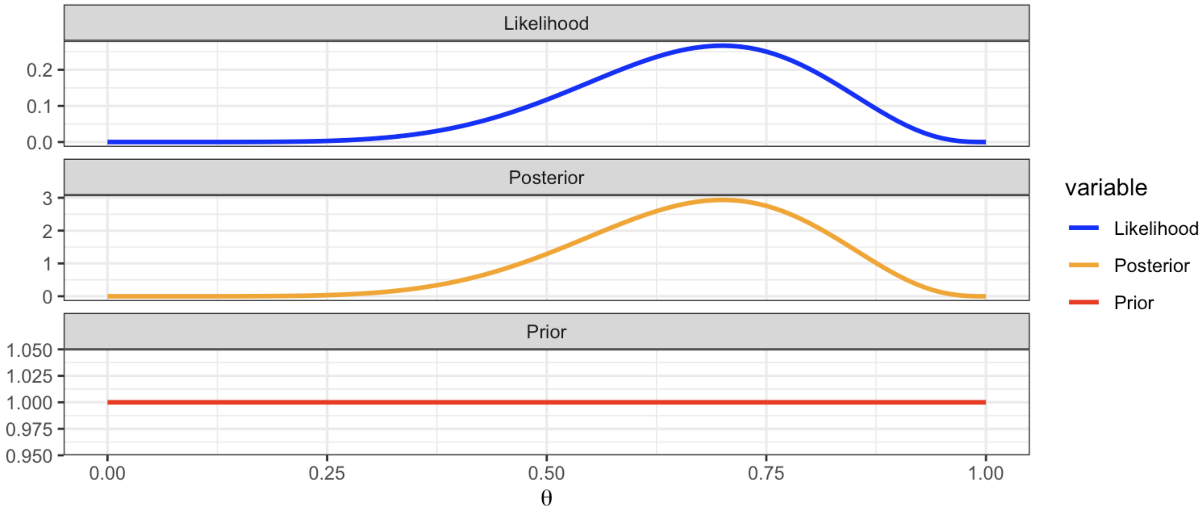
1.3)

```r
#1.3
max_index <- which.max(posterior$pst)
theta_max <- posterior$theta[max_index]
ggplot(posterior, aes(x=theta, y=pst))+geom_line(linewidth=1, color="red")+
  theme_bw()+xlab(expression(theta))+ ylab("Posterior Distribution")+
  geom_vline(xintercept=theta_max, color ='orange',linetype='dashed')+
  annotate("text", x=theta_max, y=max(posterior$pst), label=paste("theta ==",
      round(theta_max, 2)), parse=TRUE, vjust = 0, hjust = 1.1, color = 'blue')
```



1.4)

```r
#1.4
posterior$prior <- 1
posterior$lkl <- NA
for (i in 1:length(theta)){
  posterior$lkl[i] <- prod(dbinom(x=7, size=10, prob=theta[i]))
}
df.lkl_prior_pst <- melt(posterior, id=c('theta'))
df.lkl_prior_pst$variable <- ifelse(df.lkl_prior_pst$variable=="lkl",
                                    "Likelihood",
                                    ifelse(df.lkl_prior_pst$variable=='pst',
                                           'Posterior', 'Prior'))
ggplot(df.lkl_prior_pst, aes(x=theta, y=value, color=variable)) + geom_line(size=1)+
  theme_bw() + xlab(expression(theta)) + ylab("") +
  scale_x_continuous(limits = c(0,1))+
  facet_wrap(~variable, scales='free_y', ncol=1) +
  scale_color_manual(values=c("blue", "orange", "red"))
```



Note: In the question, I haven't explicitly assigned p(θ) = 0 for θ < 0 or θ > 1 as the range of values for θ is in [0, 1], hence I have ignored it in calculations for posterior and likelihood function.

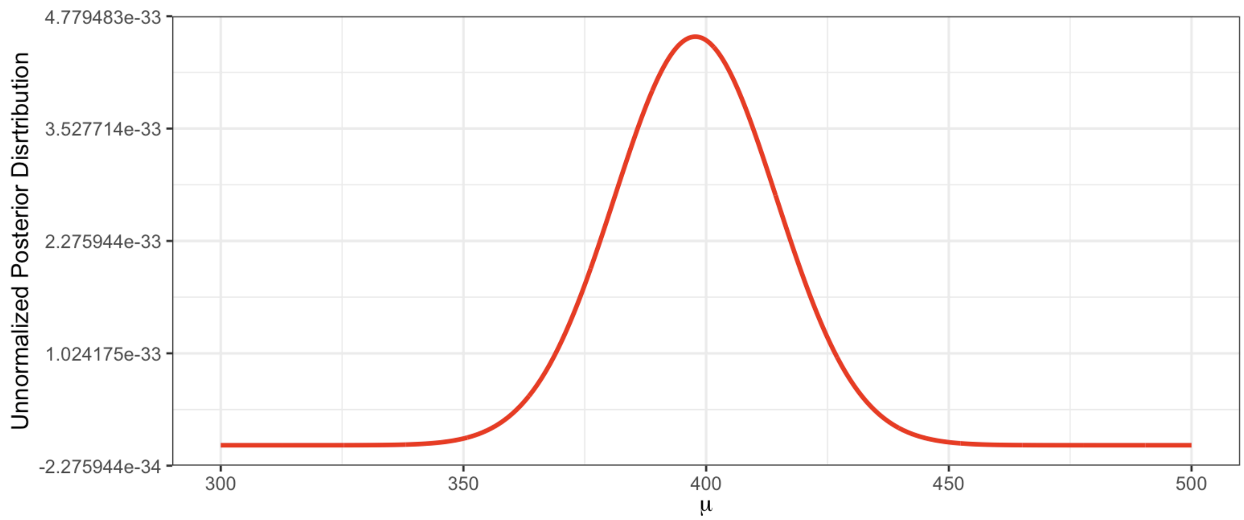
Part 2)2.1)

```
#2.1
y <- c(300, 270, 390, 450, 500, 290, 680, 450)
pst_300 <- prod(dnorm(y, 300, sd=50))*dnorm(300, mean = 250, sd = 50)
pst_900 <- prod(dnorm(y, 900, sd=50))*dnorm(900, mean = 250, sd = 50)
pst_50 <- prod(dnorm(y, 50, sd=50))*dnorm(50, mean = 250, sd = 50)
cat("Unnormalized Posterior at 300:", pst_300, "\n")
cat("Unnormalized Posterior at 900:", pst_900, "\n")
cat("Unnormalized Posterior at 50:", pst_50, "\n")
```

```
Unnormalized Posterior at 300: 1.529208e-40
> cat("Unnormalized Posterior at 900:", pst_900, "\n")
Unnormalized Posterior at 900: 3.052368e-230
> cat("Unnormalized Posterior at 50:", pst_50, "\n")
Unnormalized Posterior at 50: 1.28358e-127
```

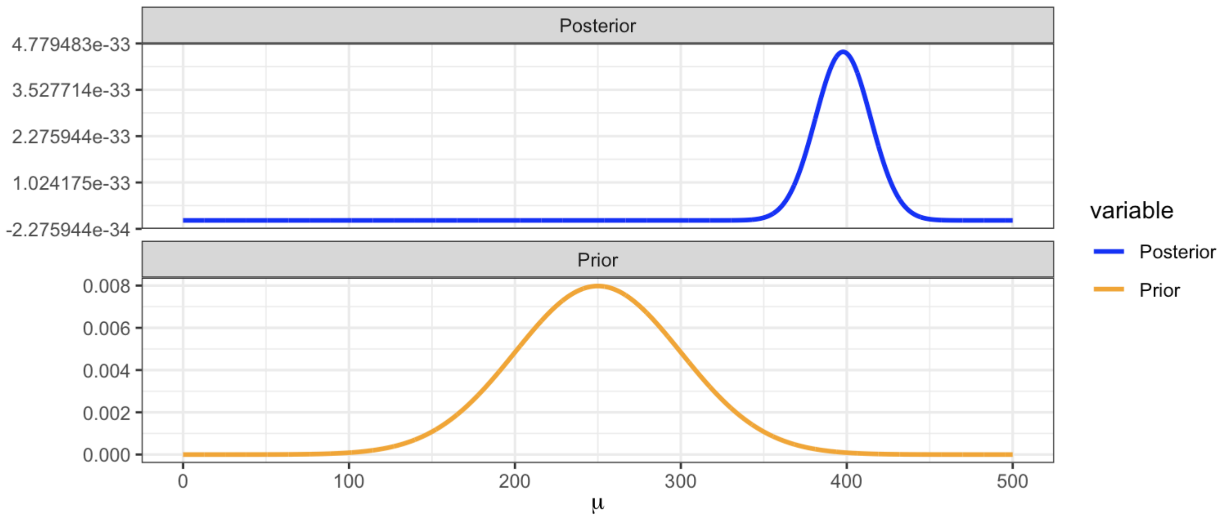
2.2)

```r
#2.2
mu <- seq(from=0, to=500, by=0.1)
posterior <- data.frame(mu=mu)
posterior$pst <- NA
posterior$prior <- NA
for (i in 1:length(mu)){
  posterior$prior[i] <- dnorm(mu[i], mean=250, sd=50)
}
for (i in 1:length(mu)){
  posterior$pst[i] <- prod(dnorm(y, mu[i], sd=50))*posterior$prior[i]
}
ggplot(posterior, aes(x=mu, y=pst))+geom_line(size=1,color='red')+
  theme_bw()+xlab(expression(mu))+ylab("Unnormalized Posterior Disrtribution")+
  scale_x_continuous(limits=c(300, 500))
```



2.3)

```r
#2.3
df.pst_prior = melt(posterior, id=c('mu'))
df.pst_prior$variable <- ifelse(df.pst_prior$variable=="prior",'Prior','Posterior')
ggplot(df.pst_prior, aes(x=mu, y=value, color=variable))+geom_line(linewidth=1)+
  theme_bw()+xlab(expression(mu))+ylab("")+
  scale_x_continuous(limits=c(0,500))+
  facet_wrap(~variable, scales='free_y',ncol=1)+
  scale_color_manual(values=c('blue','orange'))
```



Part 3)

3) Prior : $\lambda \sim \text{Gamma}(40, 2)$

Posterior : $\lambda \sim \text{Gamma}(40+k, 2+1)$

Now, this posterior becomes prior for next day

Day 1: $\lambda \sim \text{Gamma}(40, 2)$

Day 2: $\lambda \sim \text{Gamma}(65, 3)$ $k = 25$

Day 3: $\lambda \sim \text{Gamma}(85, 4)$ $k = 20$

Day 4: $\lambda \sim \text{Gamma}(108, 5)$ $k = 23$

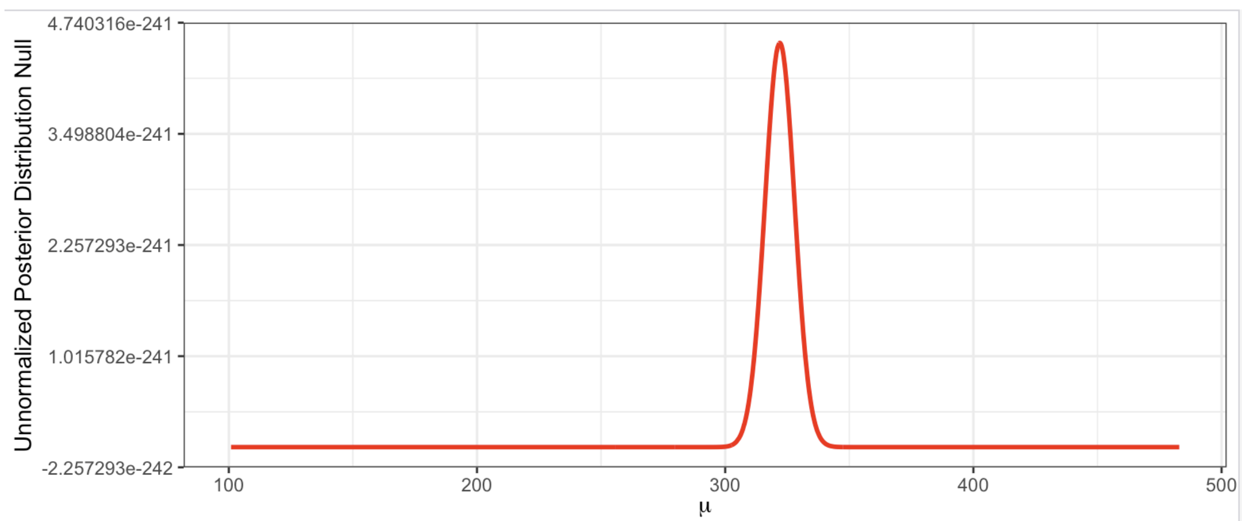
Day 5: $\underbrace{\lambda \sim \text{Gamma}(135, 6)}_{\text{Prior for Day 5}}$ $k = 27$

3.2) Number of road accidents $= \frac{135}{6} = 22.5$

Part 4)

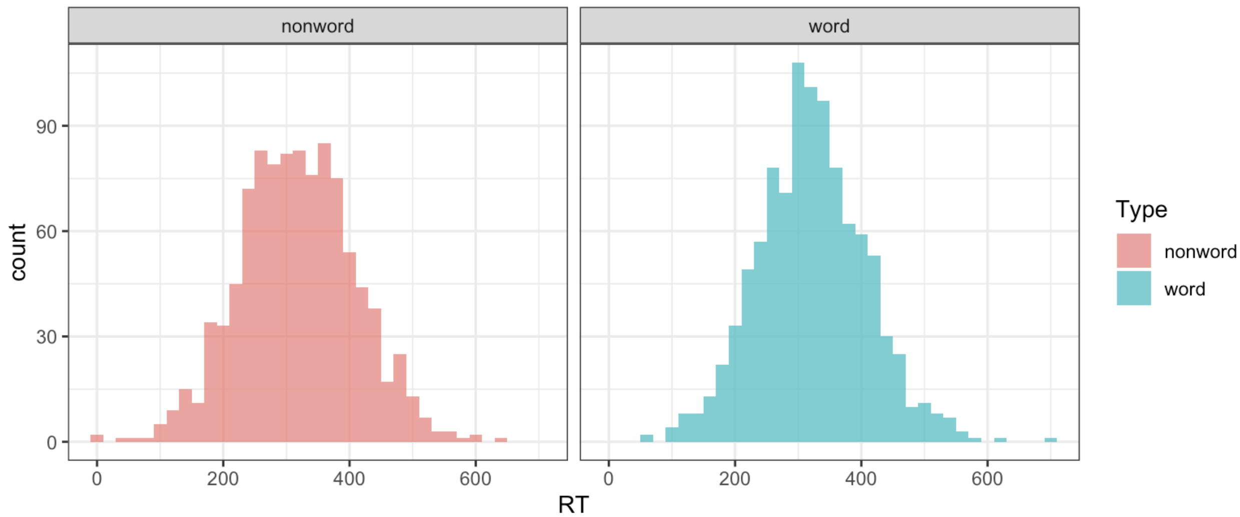
4.5.1)

```
mu <- rnorm(1000, 300, 60)
dat <- read.table(
  "https://raw.githubusercontent.com/yadavhimanshu059/CGS698C/main/notes/Module-2/rec
  sep=",",header = T)[,-1]
posterior_null <- data.frame(mu=mu)
posterior_null$prior <- NA
for(i in 1:length(mu)){
  posterior_null$prior[i] <- dnorm(mu[i], 300, 50)
}
posterior_null$pst <- NA
for(i in 1:length(mu)){
  posterior_null$pst[i] <- prod(dnorm(dat$Tw,mean=mu[i],sd=60))*
                           prod(dnorm(dat$Tnw,mean=mu[i],sd=60))*
                           posterior_null$prior[i]
}
ggplot(posterior_null, aes(x=mu, y=pst)) + geom_line(size=1,color='red')+
         theme_bw() + xlab(expression(mu)) +
         ylab("Unnormalized Posterior Distribution Null")
.
```



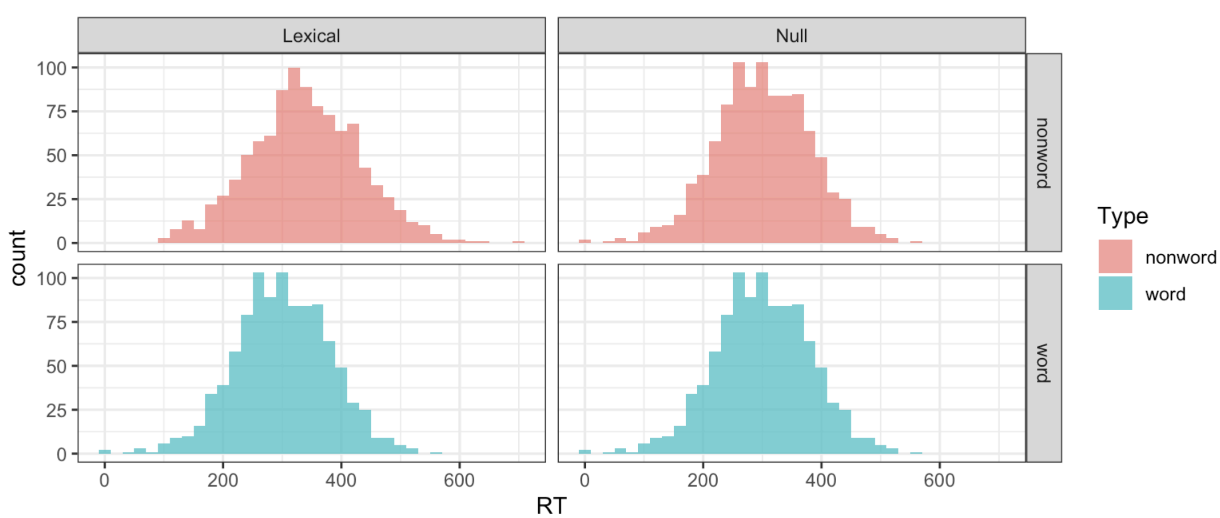
4.5.2)

```
rt_word <- rnorm(1000 ,mean=mu, sd=50)
rt_nonword <- rnorm(1000, mean = mu+delta, sd=60)
df_lexical<- data.frame(
  RT = c(rt_word, rt_nonword),
  Type = c("word", "nonword")
)
ggplot(df_lexical, aes(x=RT, fill=Type))+geom_histogram(binwidth = 20, alpha=0.7)+
  theme_bw() + facet_grid(~Type)
```



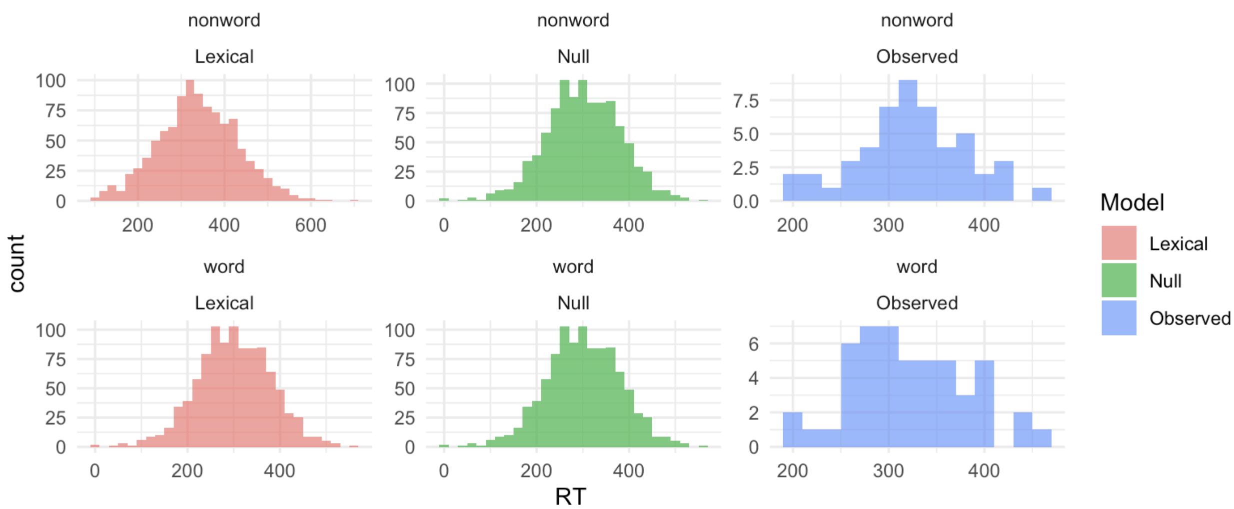
4.5.3)

```
#4.5.3
df_lexical$Model <- "Lexical"
df_null <- data.frame(
  RT = c(rt_word, rt_word),
  Type = rep(c("word", "nonword"), each=1000),
  Model = rep("Null", 2000)
)
combined_df <- rbind(df_lexical, df_null)
ggplot(combined_df, aes(x=RT, fill=Type))+geom_histogram(binwidth = 20,alpha=0.7)+
  facet_grid(Type~Model) + theme_bw()
```



4.5.4)

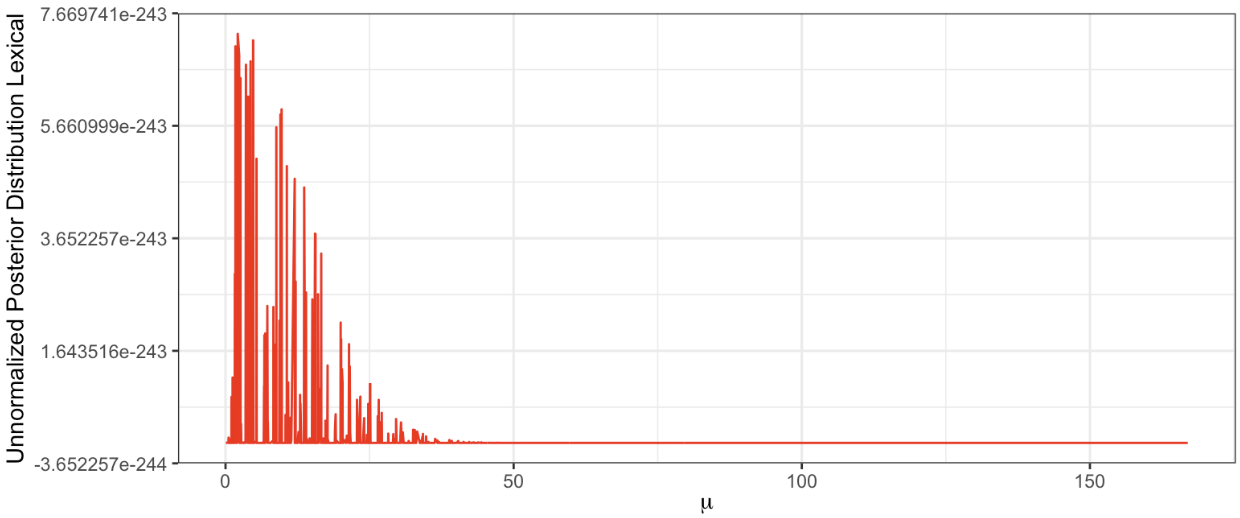
```r
#4.5.4
df_obs <- data.frame(
  RT = c(dat$Tw, dat$Tnw),
  Type = rep(c("word", "nonword"), each=nrow(dat)),
  Model = "Observed"
)
df_all <- rbind(df_null, df_lexical, df_obs)
ggplot(df_all, aes(x=RT, fill=Model))+geom_histogram(binwidth = 20, alpha=0.7)+
  theme_minimal()+facet_wrap(Type~Model, scale="free")
```



While the null and lexical models display the same recognition time for words, we can see that the lexical model tends to be more consistent with the recognition time for the non words. Hence, the lexical model appears to be more overall consistent with the data.

4.5.5)

```r
#4.5.5
library(truncnorm)
delta <- rtruncnorm(1000, a=0, b=Inf, mean=0, sd=50)
posterior_lexical <- data.frame(mu=mu, delta=delta)
posterior_lexical$pst <- NA
for (i in 1:length(mu)){
  posterior_lexical$pst[i] <- prod(dnorm(dat$Tw,mean=mu[i],sd=60))*
                              prod(dnorm(dat$Tnw,mean=mu[i]+delta[i],sd=60))*
                              dnorm(mu[i], 300, 50)*
                              dtruncnorm(delta[i],a = 0,b = Inf,mean = 0,sd = 50)
}
ggplot(posterior_lexical, aes(x=delta, y=pst)) + geom_line(color='red')+
  theme_bw() + xlab(expression(mu)) +
  ylab("Unnormalized Posterior Distribution Lexical")
```



Unnormalized Posterior Distribution Lexical
7.669741e-243
5.660999e-243
3.652257e-243
1.643516e-243
-3.652257e-244
0
50
100
150
μ