district model

Barbara Mottey

4/21/2021

```
dat <- readRDS('C:/Users/sympl/Documents/UMass/msthesis/Data/2014data.rds')


ge.shp<-readOGR("C:/Users/sympl/Documents/UMass/msthesis/GPS/GHGE71FL/GHGE71FL.shp")
```

```
## OGR data source with driver: ESRI Shapefile
## Source: "C:\Users\sympl\Documents\UMass\msthesis\GPS\GHGE71FL\GHGE71FL.shp", layer: "GHGE71FL"
## with 427 features
## It has 20 fields
```

```
bound<-readOGR("C:/Users/sympl/Documents/UMass/msthesis/GPS/sdr_subnational_boundaries_2021-03-05/shps/s
```

```
## OGR data source with driver: ESRI Shapefile
## Source: "C:\Users\sympl\Documents\UMass\msthesis\GPS\sdr_subnational_boundaries_2021-03-05\shps\sdr_s
## with 10 features
## It has 27 fields
```

```
# district boundary
dist<-readOGR("C:/Users/sympl/Documents/UMass/msthesis/GPS/Ghana_District_CORRECT/Ghana_districts_corre
```

```
## OGR data source with driver: ESRI Shapefile
## Source: "C:\Users\sympl\Documents\UMass\msthesis\GPS\Ghana_District_CORRECT\Ghana_districts_correct.s
## with 110 features
## It has 18 fields
```

```
#From RA spatial points
dist3<- readOGR("C:/Users/sympl/Documents/UMass/msthesis/GPS/Ghana_Dist_DHS_Join/GPS_Points_Districts.sh
```

```
## OGR data source with driver: ESRI Shapefile
## Source: "C:\Users\sympl\Documents\UMass\msthesis\GPS\Ghana_Dist_DHS_Join\GPS_Points_Districts.shp", l
## with 427 features
## It has 41 fields
## Integer64 fields read as strings:  OBJECTID Join_Count TARGET_FID Index Household
```

```
#plot(dist)
#points(dist3, pch=".", col="red")
```

DISTRICT

```
nb1 <- poly2nb(dist, row.names = dist@data$DIST_NAME)#for calculating neighbors

nb2INLA("map.adj", nb1)
g1 <- inla.read.graph(filename = "map.adj") #neighbour of each area for modeling


##define stuctured and unstructured spatial re vectors
dist$re_u <- 1:nrow(dist@data)
dist$re_v <- 1:nrow(dist@data)

#rename data
trial1<-dat
trial1$cluster.no<- as.numeric(trial1$cluster.no)

#join dist and dist3
k2 <- dist@data %>% select(DIST_NAME, re_v, re_u)
dist3@data<-dist3@data %>% right_join(k2, by= c("DIST_NAME"="DIST_NAME"))


trial1<-trial1 %>% right_join(dist3@data, by= c("cluster.no"="DHSCLUST"))
```

Model

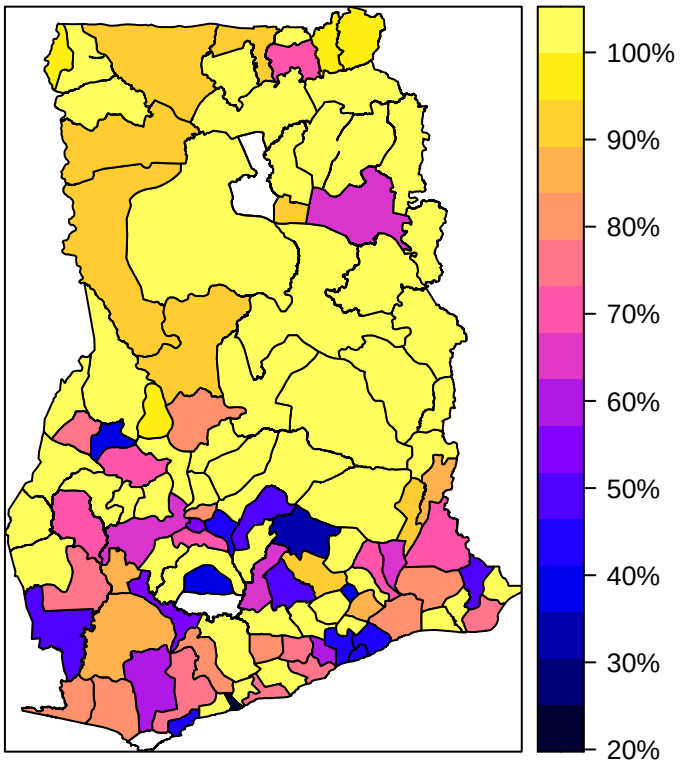
```
##spatial

formula <- c_weight ~ fuel_bin+gender+education+ w_age+ marital_s+wealth+bmi+residence+ f(re_u, model =

res2 <- inla(formula, family = "gaussian", data = trial1,  control.predictor = list(compute = TRUE))

#summary(res2)


#no spatial
formula <- c_weight ~ fuel_bin+gender+education+ w_age+ marital_s+wealth+bmi+residence

# inla
res2a <- inla(formula, family = "gaussian", data = trial1,  control.predictor = list(compute = TRUE))
```

maps

```
#map of solid fuel % per district
prop.fuel<- trial1 %>% filter(fuel_bin != is.na(fuel_bin)) %>%
                       group_by(DIST_NAME, fuel_bin) %>%
                       summarise(n=n()) %>%
                       mutate(solid.prop= n/sum(n)) %>%
                       filter(fuel_bin ==1)
```

```
## `summarise()` regrouping output by 'DIST_NAME' (override with `.groups` argument)
```

```
dist@data<-dist@data %>% left_join(prop.fuel, by= c("DIST_NAME" = "DIST_NAME"))
spplot(dist, "solid.prop" , colorkey = list(labels = list( labels = c("20%", "30%","40%","50%","60%","70
```

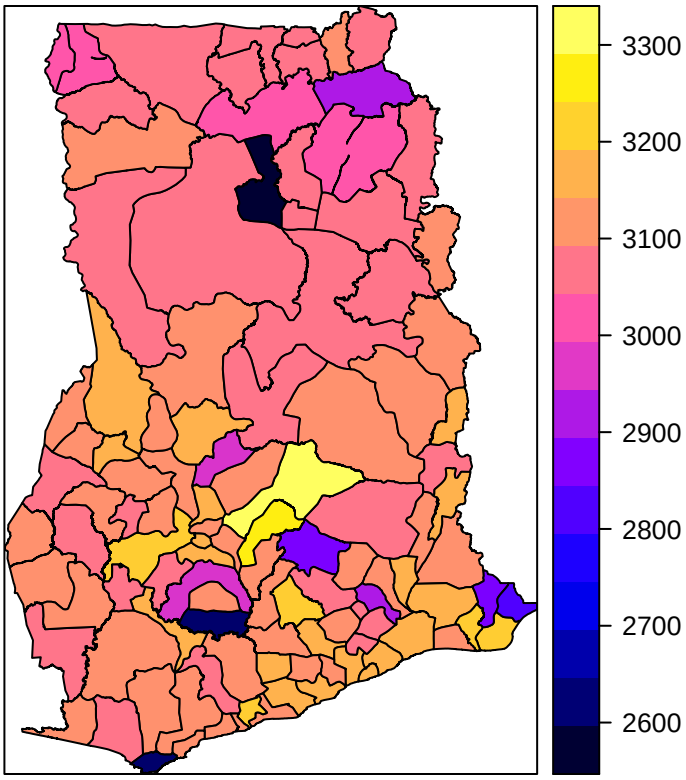


solid fuel percentage by district

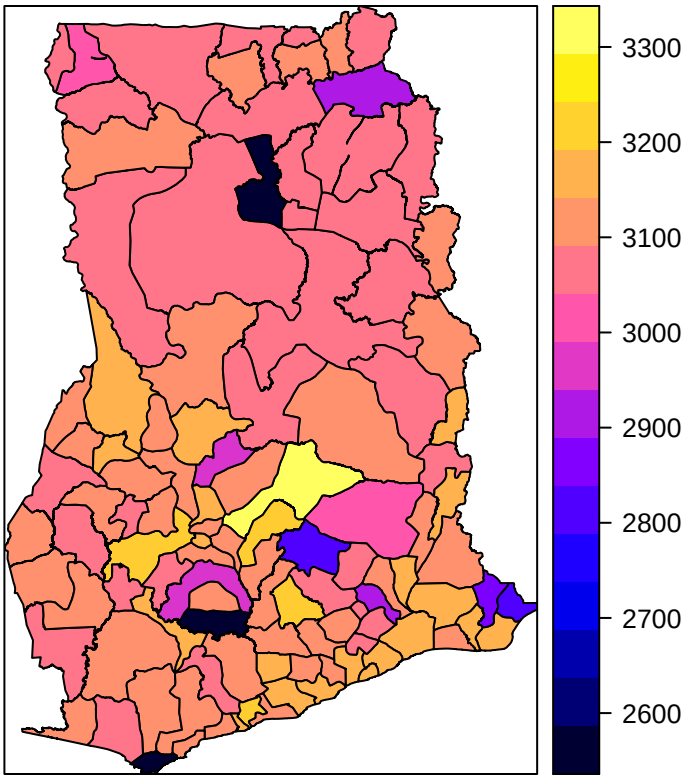
```
#map of posterior bw (spatial vs linear)
trial1$RR <- res2$summary.fitted.values[,1]
trial1$RR1 <- res2a$summary.fitted.values[,1]
new3 <- trial1 %>% group_by(DIST_NAME) %>%
  summarize(meanbwdis = mean(RR, na.rm=T), meanbwdis.lm = mean(RR1, na.rm=T))
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
dist@data<-dist@data %>% left_join(new3, by= c("DIST_NAME" = "DIST_NAME"))
spplot(dist, "meanbwdis")
```



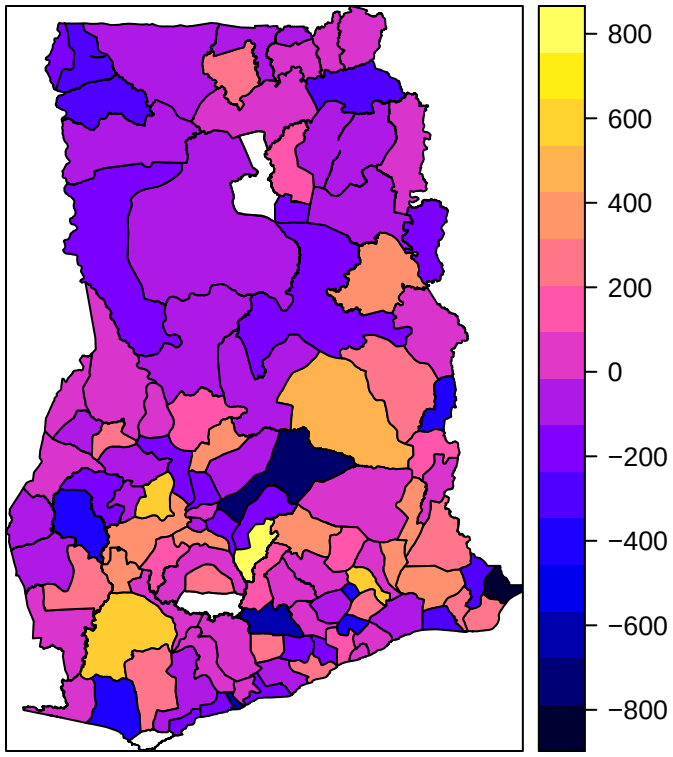

```
spplot(dist, "meanbwdis.lm")
```




```
#map of residuals
residuals2=  trial1$c_weight-res2$summary.fitted.values[,1] #spatial
trial1$residuals2<-residuals2

residuals2a =  trial1$c_weight-res2a$summary.fitted.values[,1] #linear
trial1$residuals2a<-residuals2a

new3 <- trial1 %>% group_by(DIST_NAME) %>%
  summarize(meanres2 = mean(residuals2, na.rm=T), meanres2a = mean(residuals2a, na.rm=T))
```

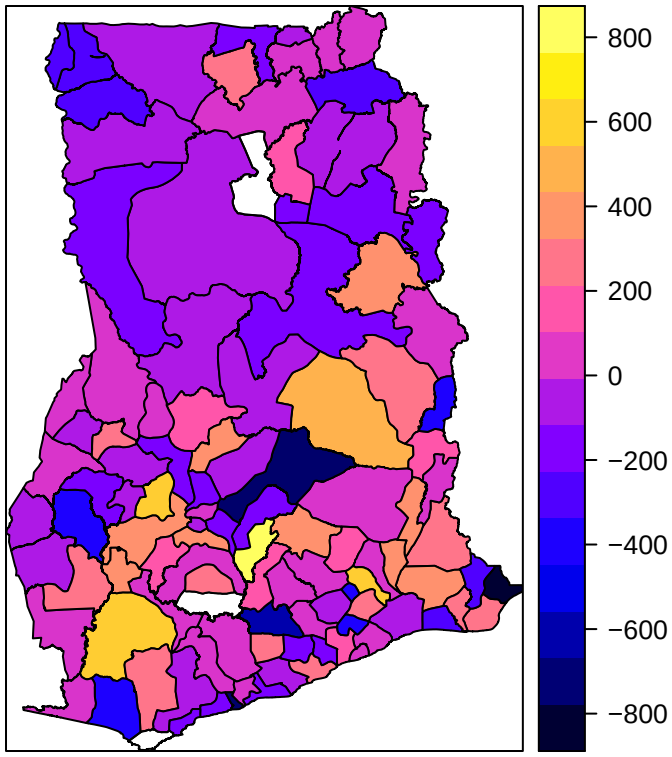
```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
dist@data<-dist@data %>% left_join(new3, by= c("DIST_NAME" = "DIST_NAME"))


#map of CAR
new3 <- data.frame(res2$summary.random$re_u[,1:2], res2$summary.random$re_v[,2])
colnames(new3)<- c("DIST_NAME" ,"CAR" , "IID")
new3$iidcar <- new3[,2]+new3[,3]
new3$DIST_NAME <- as.character(new3$DIST_NAME)
#dist@data$DIST_NAME <- as.character(dist@data$DIST_NAME)
dist@data<-dist@data %>% left_join(new3, by= c("DIST_NAME" = "DIST_NAME"))

spplot(dist, "meanres2", xlab="spatial model residuals")
```



spatial model residuals

```
spplot(dist,"meanres2a" , xlab= "linear model residuals")
```



linear model residuals

```
#spplot(dist, "CAR", xlab= "structured spatial effect")
#spplot(dist, "IID", xlab= "unstructured spatial effect")
#spplot(dist, "iidcar", xlab= "total spatial effect")
```

moran test

```
#moran test for spatial effect
moran.test(na.omit(dist@data$meanres2), nb2listw(nb1), 110)
```

```
##
##  Moran I test under randomisation
##
## data:  na.omit(dist@data$meanres2)
## weights: nb2listw(nb1)
## omitted: 3, 60, 97
##
## Moran I statistic standard deviate = 0.097405, p-value = 0.4612
## alternative hypothesis: greater
## sample estimates:
## Moran I statistic       Expectation          Variance
##      -0.003327947      -0.009433962       0.003929668
```

```
#moran test for linear
moran.test(na.omit(dist@data$meanres2a), nb2listw(nb1), 110)
```

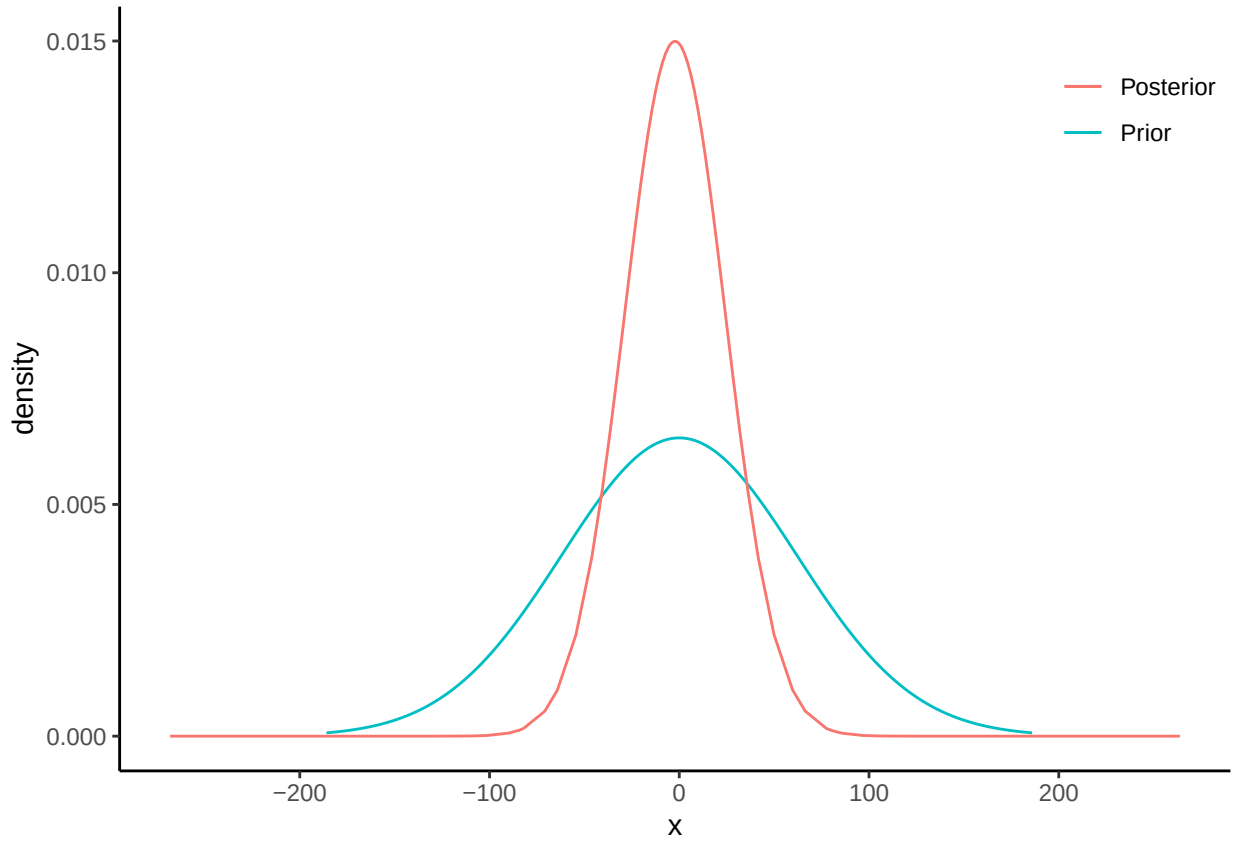
```
##
##  Moran I test under randomisation
##
## data:  na.omit(dist@data$meanres2a)
## weights: nb2listw(nb1)
## omitted: 3, 60, 97
##
## Moran I statistic standard deviate = 0.26299, p-value = 0.3963
## alternative hypothesis: greater
## sample estimates:
## Moran I statistic       Expectation          Variance
##       0.007057526      -0.009433962       0.003932299
```

density plots fixed effect

```
(SD<-(1/sqrt(0.001) * qnorm(0.975,0,1)))
```

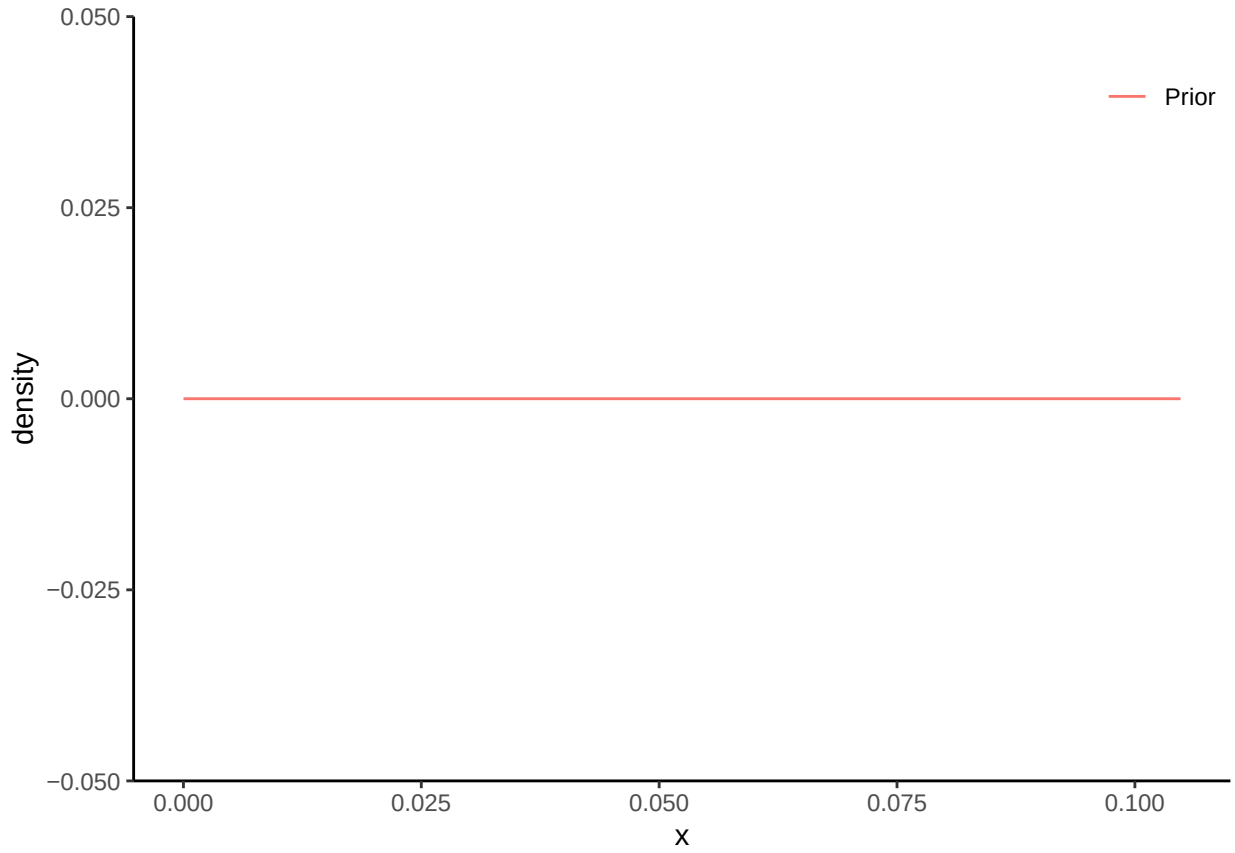
```
## [1] 61.9795
```

```
prior <- data.frame(x=seq(-3*SD,3*SD,len=150))

prior$density <- dnorm(prior$x,0,SD)
post <- data.frame(res2$marginals.fixed[[2]])


ggplot(prior, aes(y=density, x=x)) + geom_line(aes(color='Prior')) +
       geom_line(data=post, aes(y=y, x=x, color='Posterior')) +
       scale_color_discrete('')+ #labs(title = "Fuel type density", y="y")+
       theme_classic()+theme(legend.position=c(1,1), legend.justification=c(1,1))
```

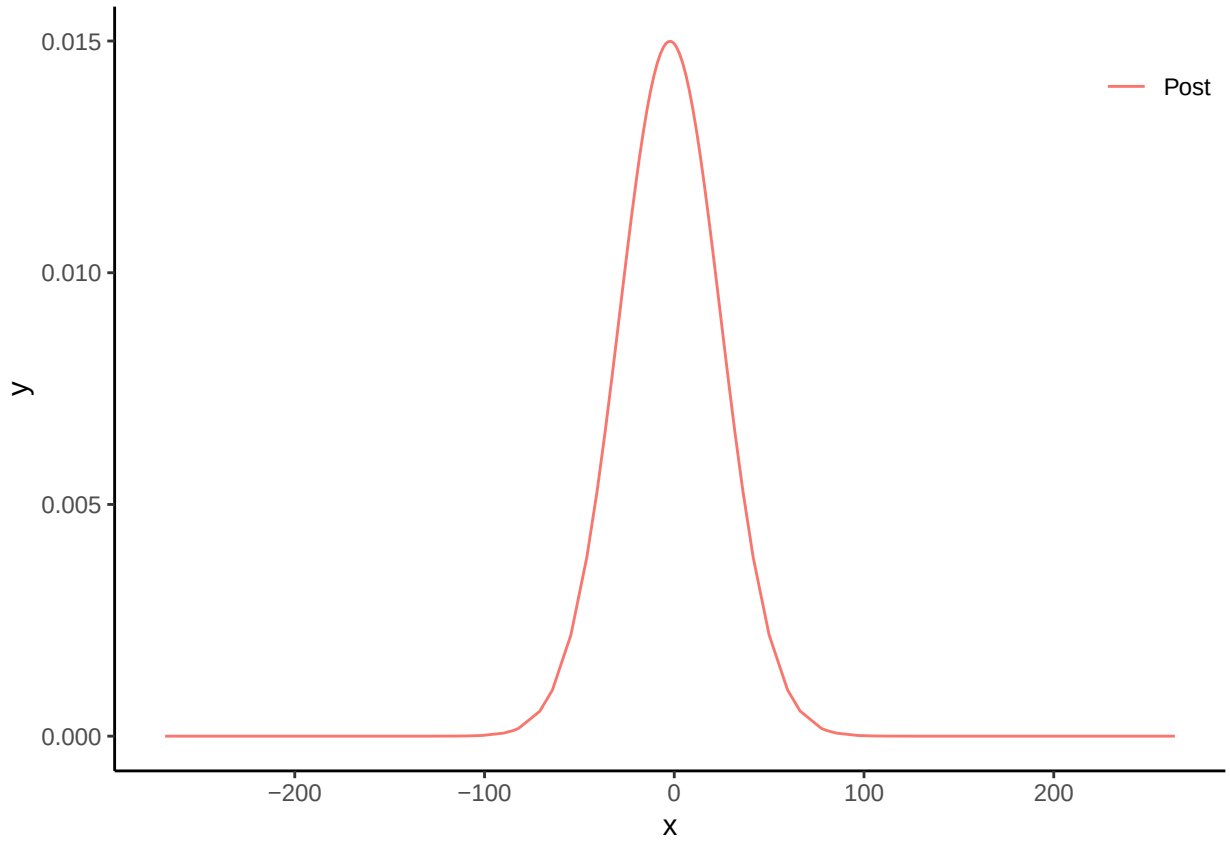


density plots random effect

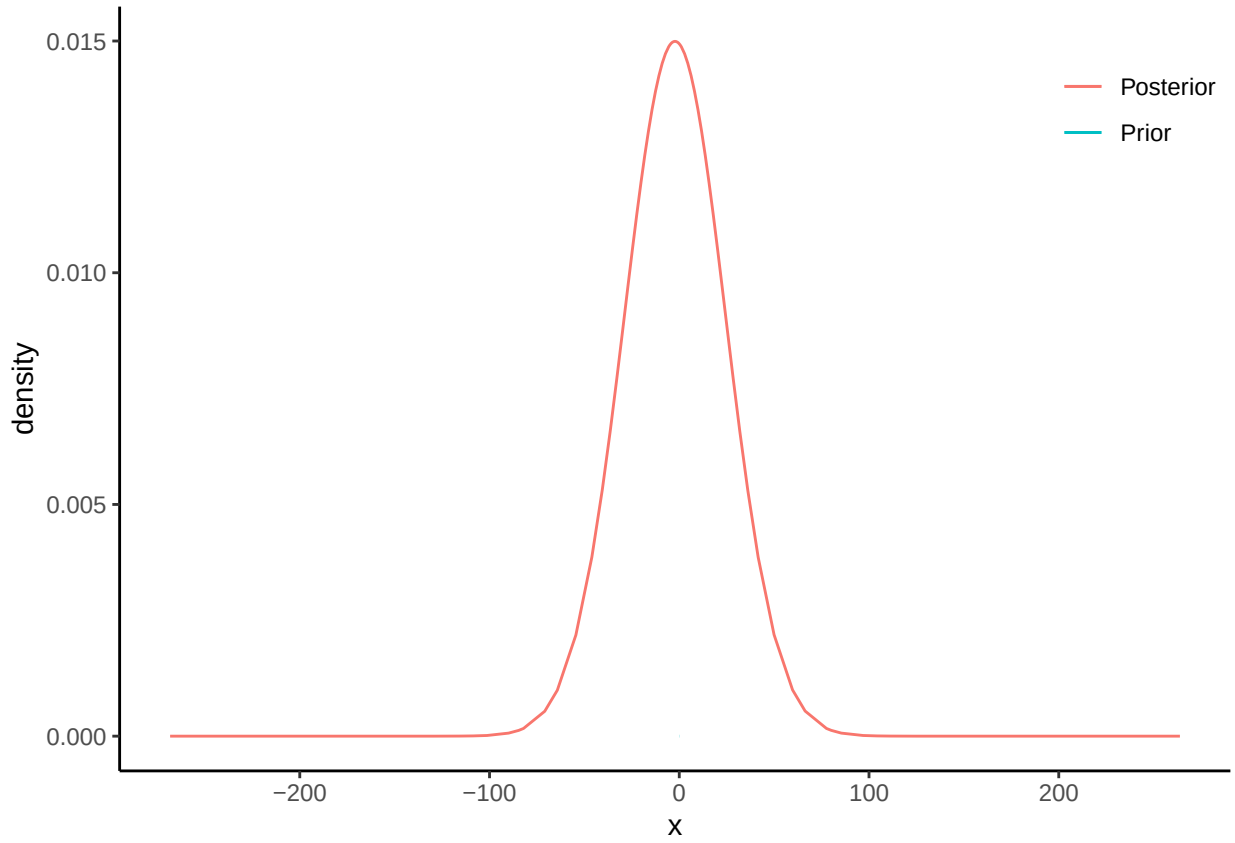
```
library(invgamma)
#inverse-gamma (a=1, b=20000)

#prior
prior <- data.frame(x=rinvgamma(150, shape=1, scale=20000))
prior$density <- dinvgamma(prior$x,1,20000)

#sigma^2 y
post1 <- data.frame(res2$marginals.hyperpar[[1]])
#prior plot
ggplot(prior, aes(y=density, x=x)) + geom_line(aes(color='Prior')) +
       #geom_line(data=post, aes(y=y, x=x, color='Posterior')) +
       scale_color_discrete('')+
       theme_classic()+theme(legend.position=c(1,1), legend.justification=c(1,1))
```



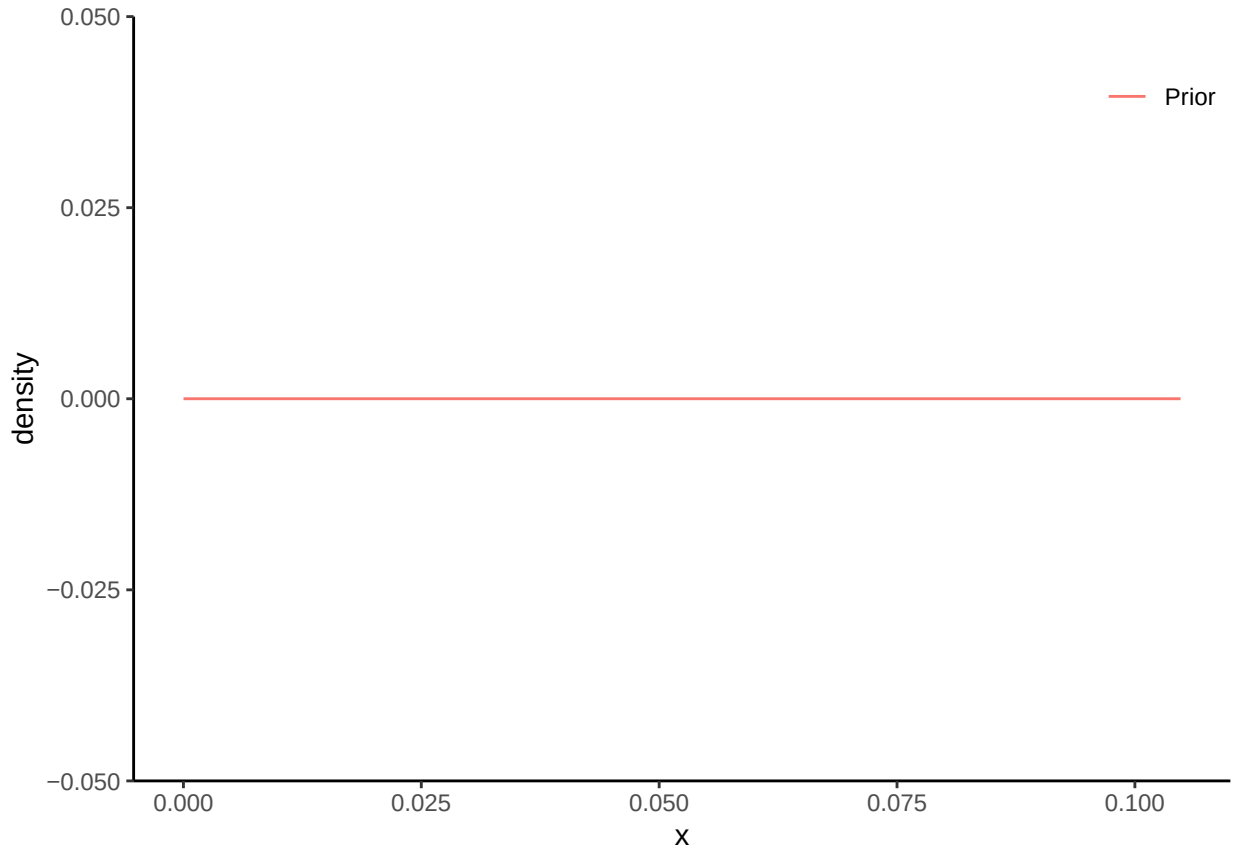

```
## posterior
ggplot(post, aes(y=y, x=x)) + geom_line(aes(color='Post')) +
       scale_color_discrete('')+
       theme_classic()+theme(legend.position=c(1,1), legend.justification=c(1,1))
```



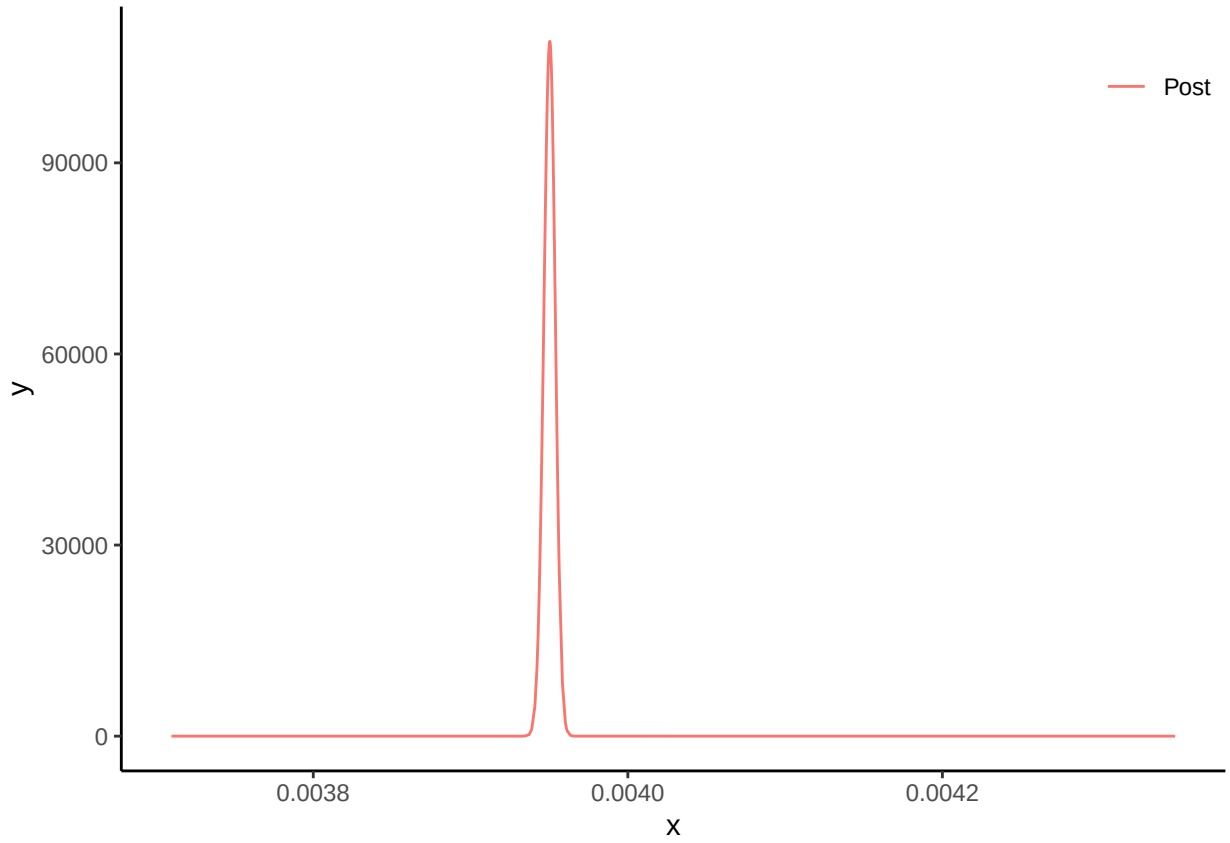

```
#both
ggplot(prior, aes(y=density, x=x)) + geom_line(aes(color='Prior')) +
       geom_line(data=post, aes(y=y, x=x, color='Posterior')) +
       scale_color_discrete('')+
       theme_classic()+theme(legend.position=c(1,1), legend.justification=c(1,1))
```



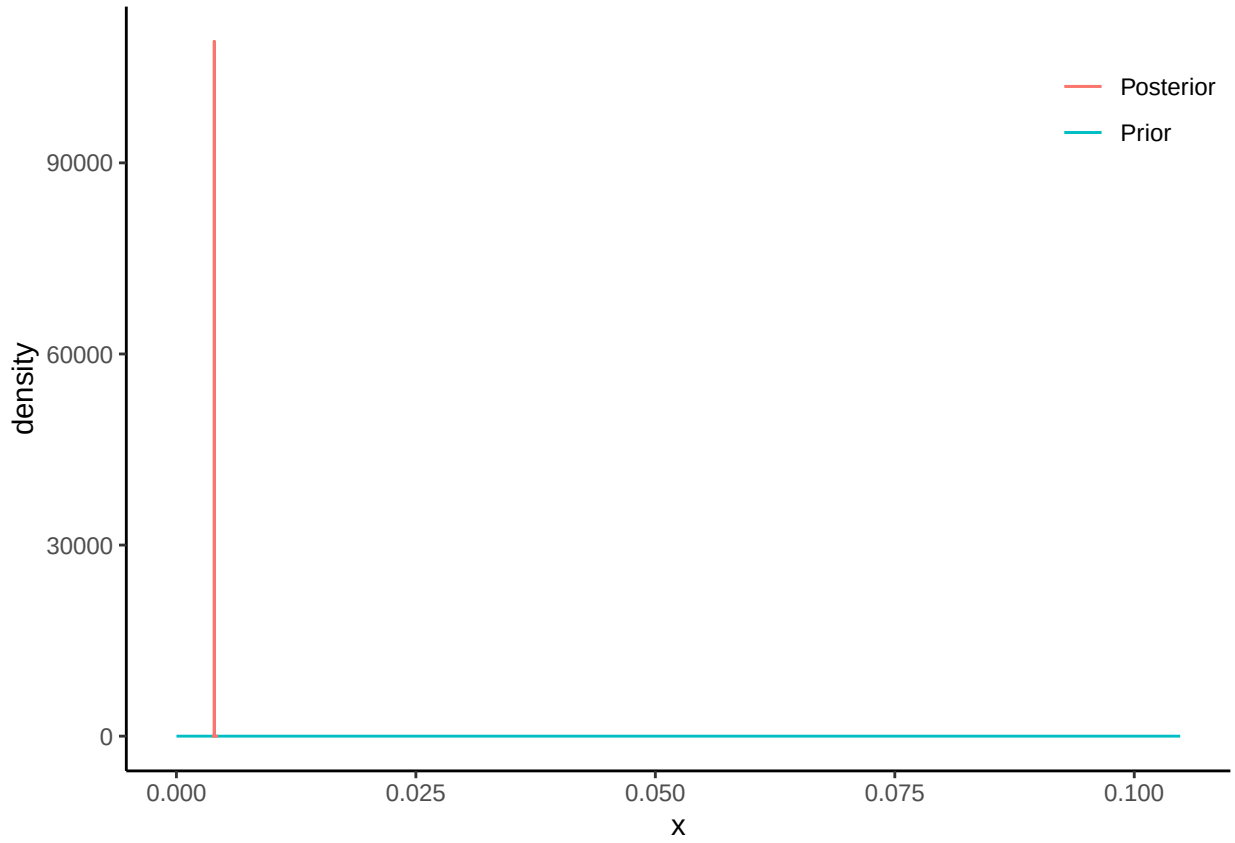

```
#tau (CAR)
post1 <- data.frame(res2$marginals.hyperpar[[2]])
#prior plot
ggplot(prior, aes(y=density, x=x)) + geom_line(aes(color='Prior')) +
       #geom_line(data=post, aes(y=y, x=x, color='Posterior')) +
       scale_color_discrete('')+
       theme_classic()+theme(legend.position=c(1,1), legend.justification=c(1,1))
```



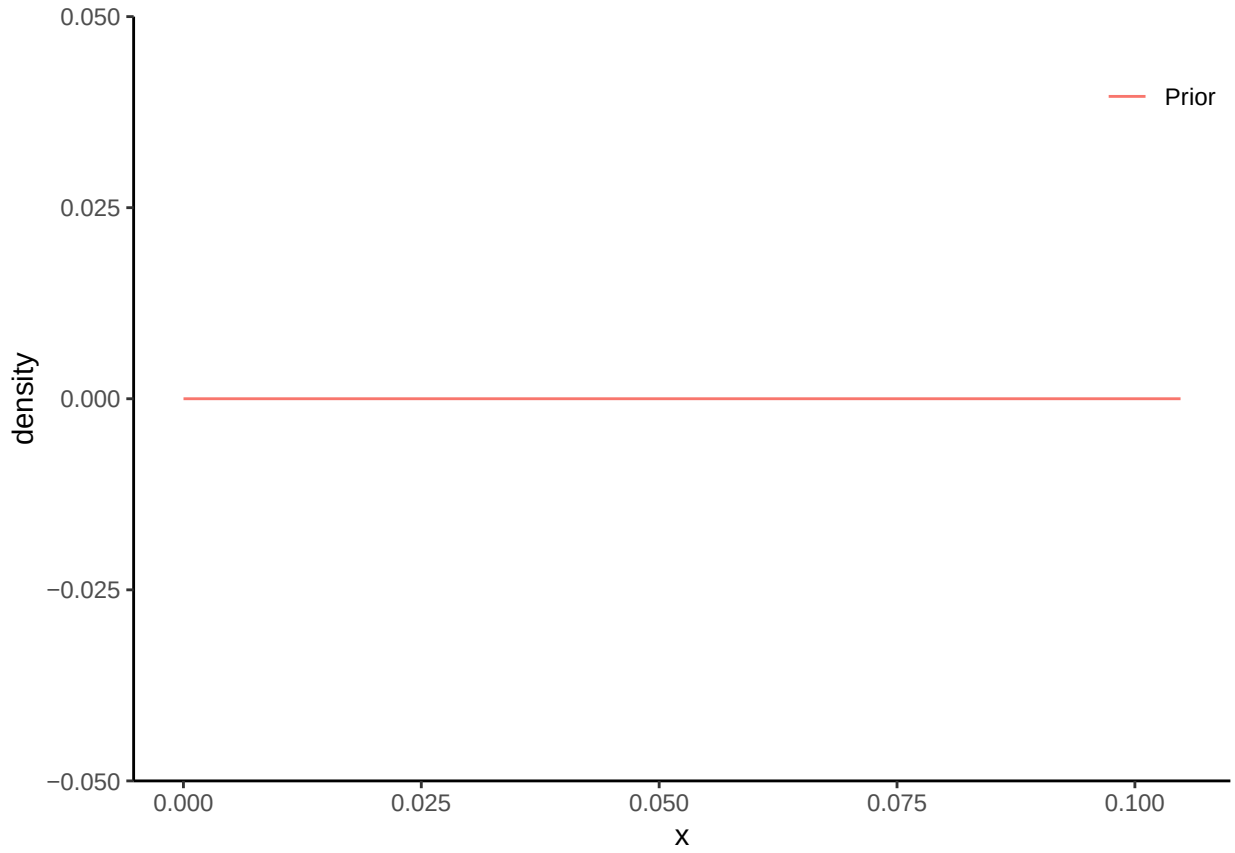

```
## posterior
ggplot(post1, aes(y=y, x=x)) + geom_line(aes(color='Post')) +
       scale_color_discrete('')+
       theme_classic()+theme(legend.position=c(1,1), legend.justification=c(1,1))
```



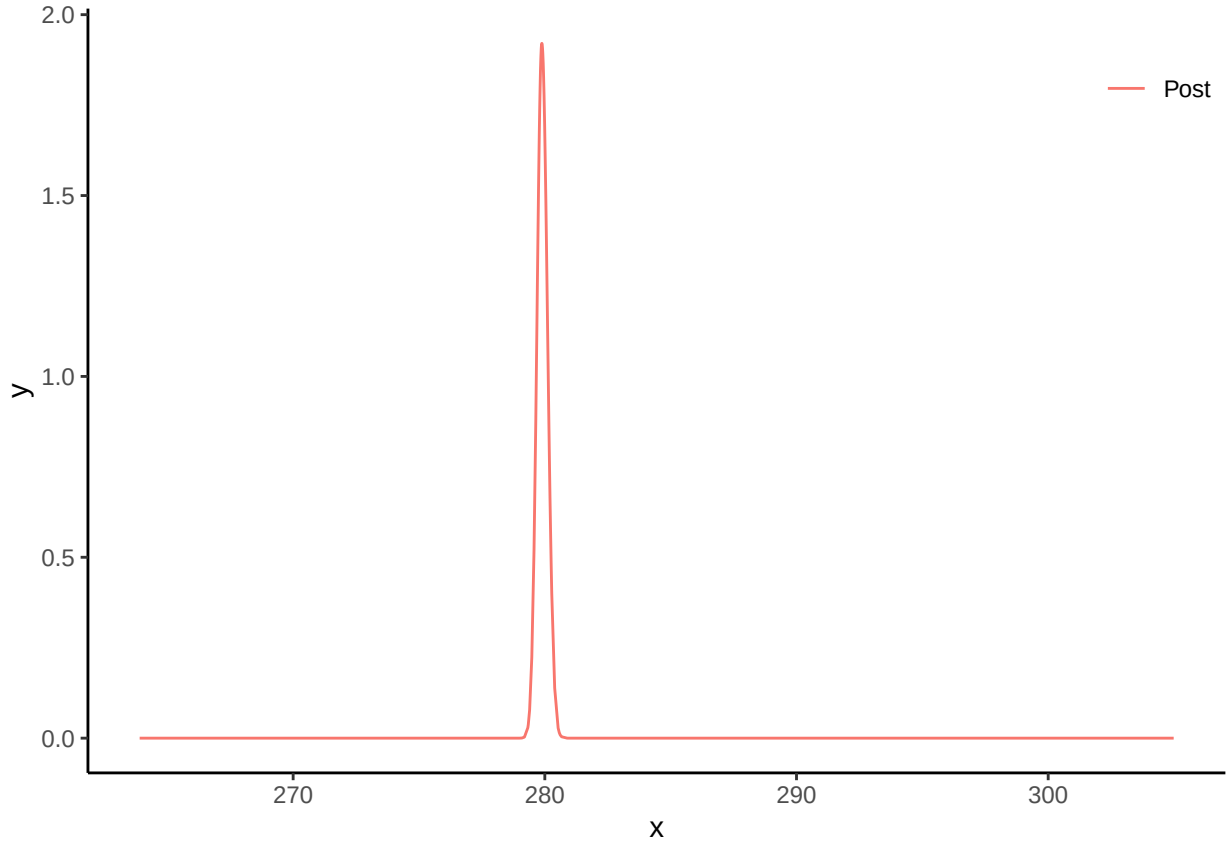

```
#both
ggplot(prior, aes(y=density, x=x)) + geom_line(aes(color='Prior')) +
       geom_line(data=post1, aes(y=y, x=x, color='Posterior')) +
       scale_color_discrete('')+
       theme_classic()+theme(legend.position=c(1,1), legend.justification=c(1,1))
```



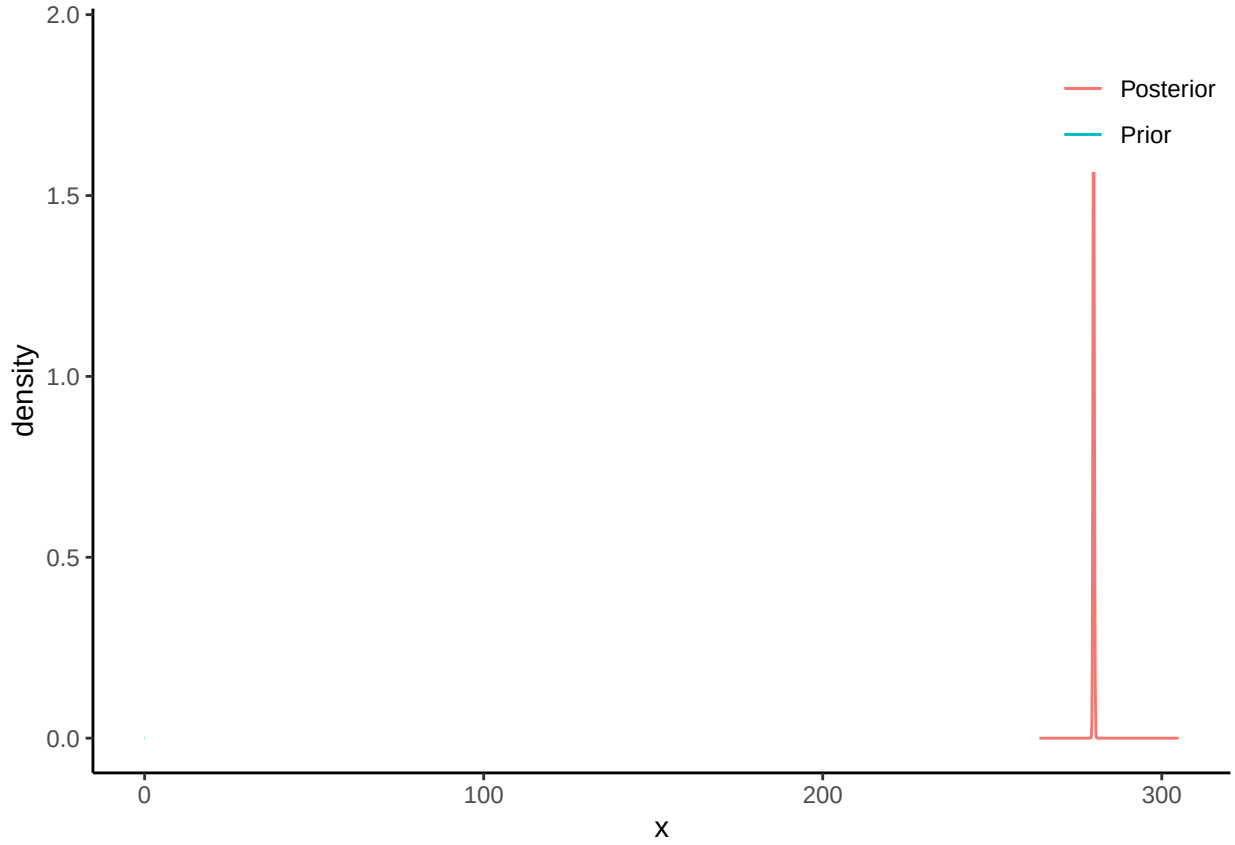

```
## sigma^2 gamma (iid)
post2 <- data.frame(res2$marginals.hyperpar[[3]])
#prior plot
ggplot(prior, aes(y=density, x=x)) + geom_line(aes(color='Prior')) +
       #geom_line(data=post, aes(y=y, x=x, color='Posterior')) +
       scale_color_discrete('')+
       theme_classic()+theme(legend.position=c(1,1), legend.justification=c(1,1))
```




```
## posterior
ggplot(post2, aes(y=y, x=x)) + geom_line(aes(color='Post')) +
       scale_color_discrete('')+
       theme_classic()+theme(legend.position=c(1,1), legend.justification=c(1,1))
```




```
#both
ggplot(prior, aes(y=density, x=x)) + geom_line(aes(color='Prior')) +
       geom_line(data=post2, aes(y=y, x=x, color='Posterior')) +
       scale_color_discrete('')+
       theme_classic()+theme(legend.position=c(1,1), legend.justification=c(1,1))
```


2.0
1.5
1.0
0.5
0.0
density
0
100
200
300
x
Posterior
Prior