

# Exam\_20221020\_Ali

Mohamed Ali - Mohal954

2023-05-23

## Question 1

D

computes the log posterior distribution

```
logpost<- function(x,N,n,a,b,theta){
  loglike<-sum(x)*log(theta) + (n*N-sum(x))*log(1-theta)
  logprior<- (a-1)*log(theta) + (b-1)*log(1-theta)
  return(-(loglike+logprior))
}

x<-c(13,8,11,7)
N=20
a=2
b=3
n=1
theta=seq(0.01,.99,.001)
theta_val<-c()
for (i in 1:length(theta)){
  theta_val[i]<-logpost(x,N,n,a,b,theta[i])
}

plot_df<- data.frame(x=theta_val)

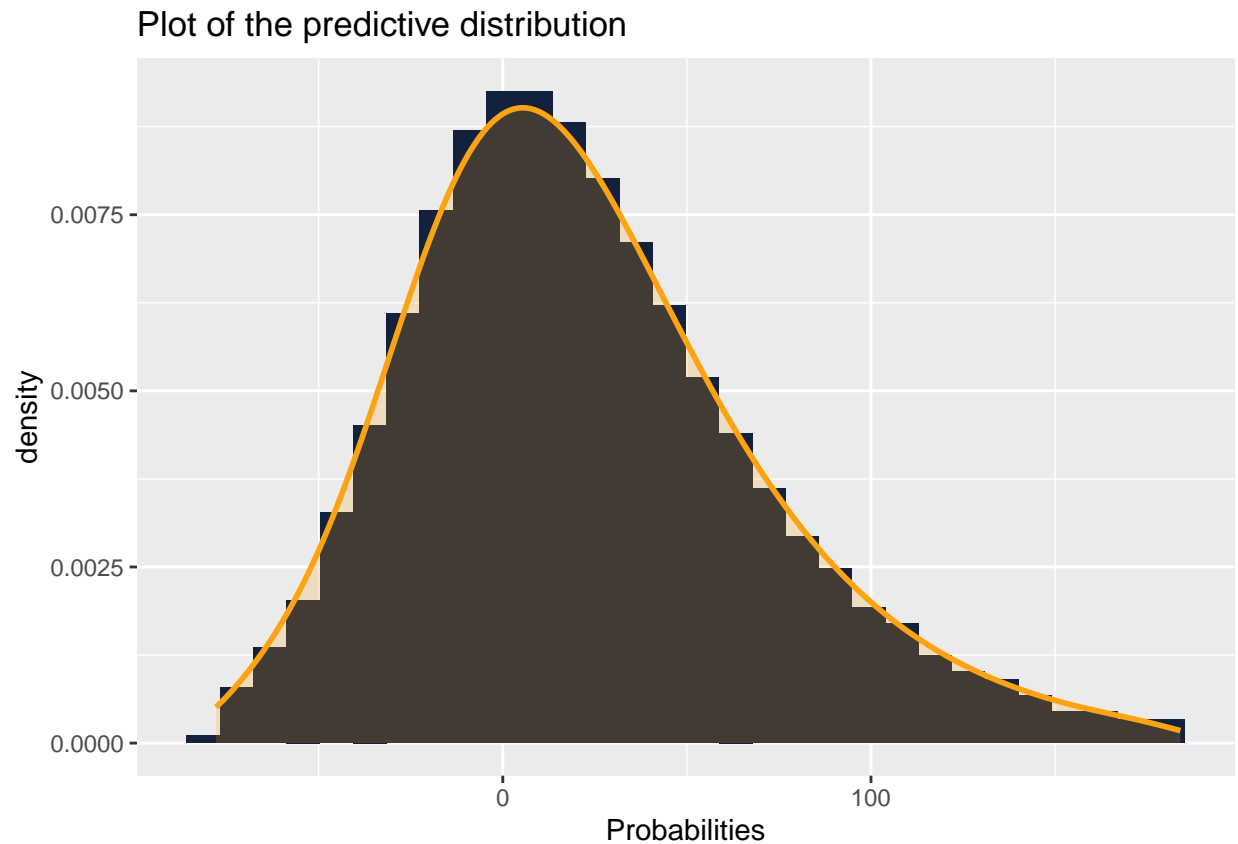
ggplot(plot_df,aes(x = x)) +
  geom_histogram(aes(y=..density..),linetype=1,fill='#14213D')+
  geom_density(alpha=.2,color="#FCA311",size=1,fill="#FCA311")+
  labs(x = 'Probabilities', y = 'density',
       title = 'Plot of the predictive distribution')
```

```
## Warning: Using 'size' aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use 'linewidth' instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```

```
## Warning: The dot-dot notation ('..density..') was deprecated in ggplot2 3.4.0.
## i Please use 'after_stat(density)' instead.
```

```
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```

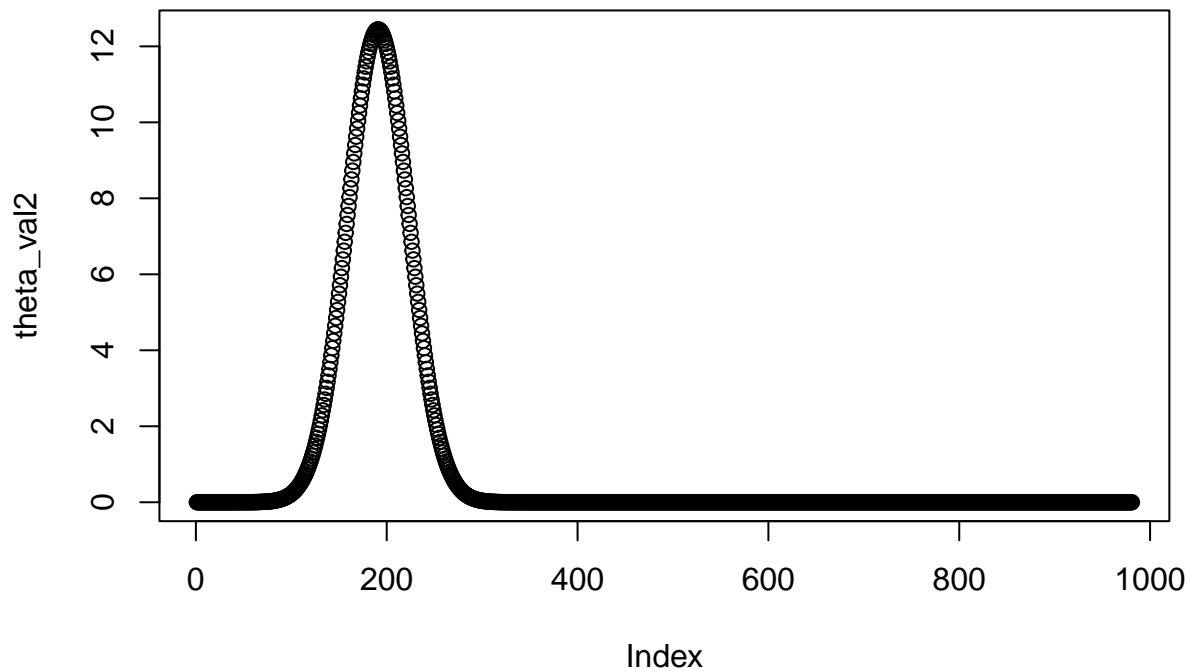


## E

You can also embed plots, for example:

```
initVal=0.5
OptimRes <- optim(initVal,
                  logpost,
                  gr=NULL,
                  N=N,
                  x=x,
                  a=a,
                  b=b,
                  n=n,
                  method=c("L-BFGS-B"),
                  control=list(fnscale=-1),
                  lower = .2,
                  hessian=TRUE)
```

```
theta_val2<-dnorm(theta,mean=OptimRes$par,sd=sqrt(1/OptimRes$hessian))
plot(theta_val2)
```



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

## Question 2

A

Plot the posterior distribution

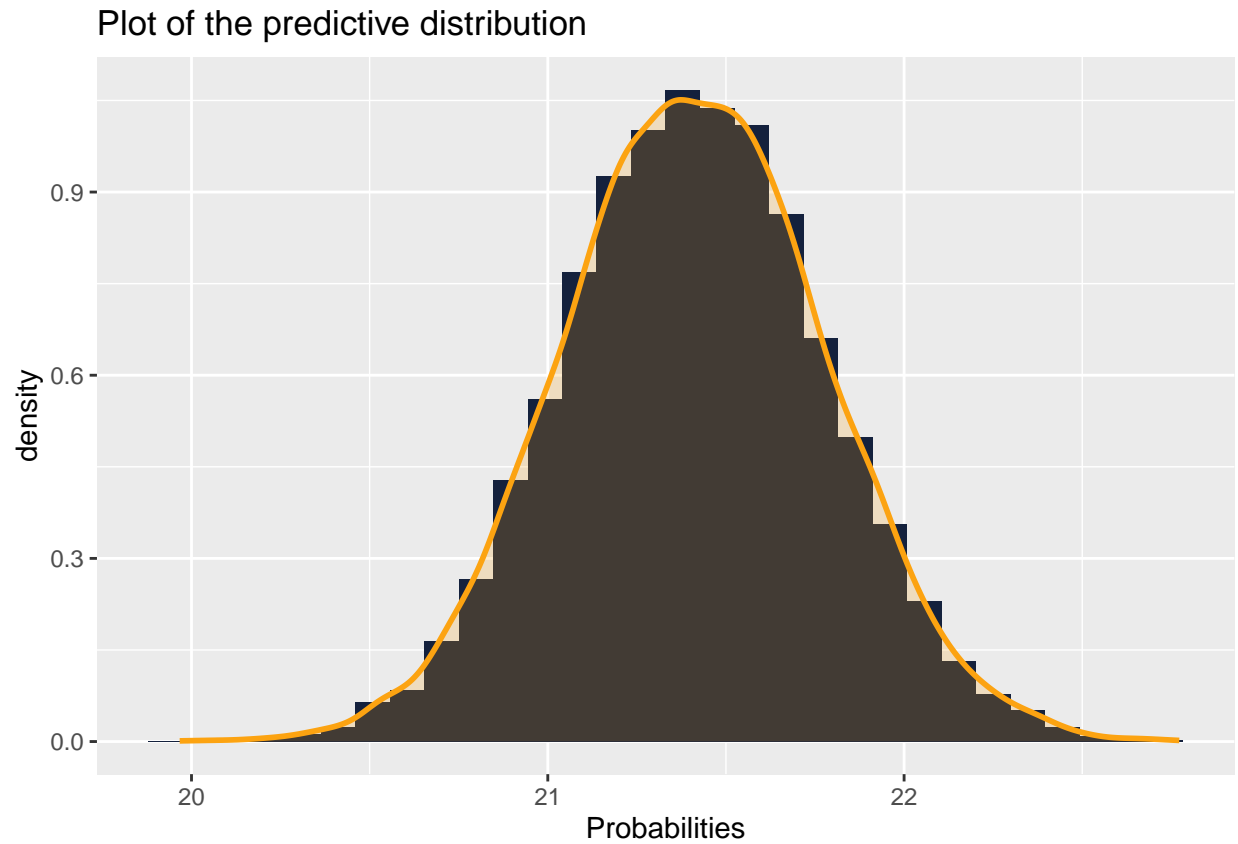
```
a=sum(Wolves$y)+40
b=length(Wolves$y)+2
plt_df<-data.frame(x=rgamma(10000,a,b))
paste0("the posterior probability is smaller than 21 = ",mean((rgamma(10000,a,b))<21))
```

```
## [1] "the posterior probability is smaller than 21 = 0.1269"
```

```
ggplot(plt_df,aes(x = x)) +
  geom_histogram(aes(y=..density..),linetype=1,fill='#14213D')+
  geom_density(alpha=.2,color="#FCA311",size=1,fill="#FCA311")+
```

```
labs(x = 'Probabilities', y = 'density',
      title = 'Plot of the predictive distribution')
```

## 'stat\_bin()' using 'bins = 30'. Pick better value with 'binwidth'.



B

```
###
reg_a<- subset(Wolves,Wolves==1)
reg_b<- subset(Wolves,Wolves==0)

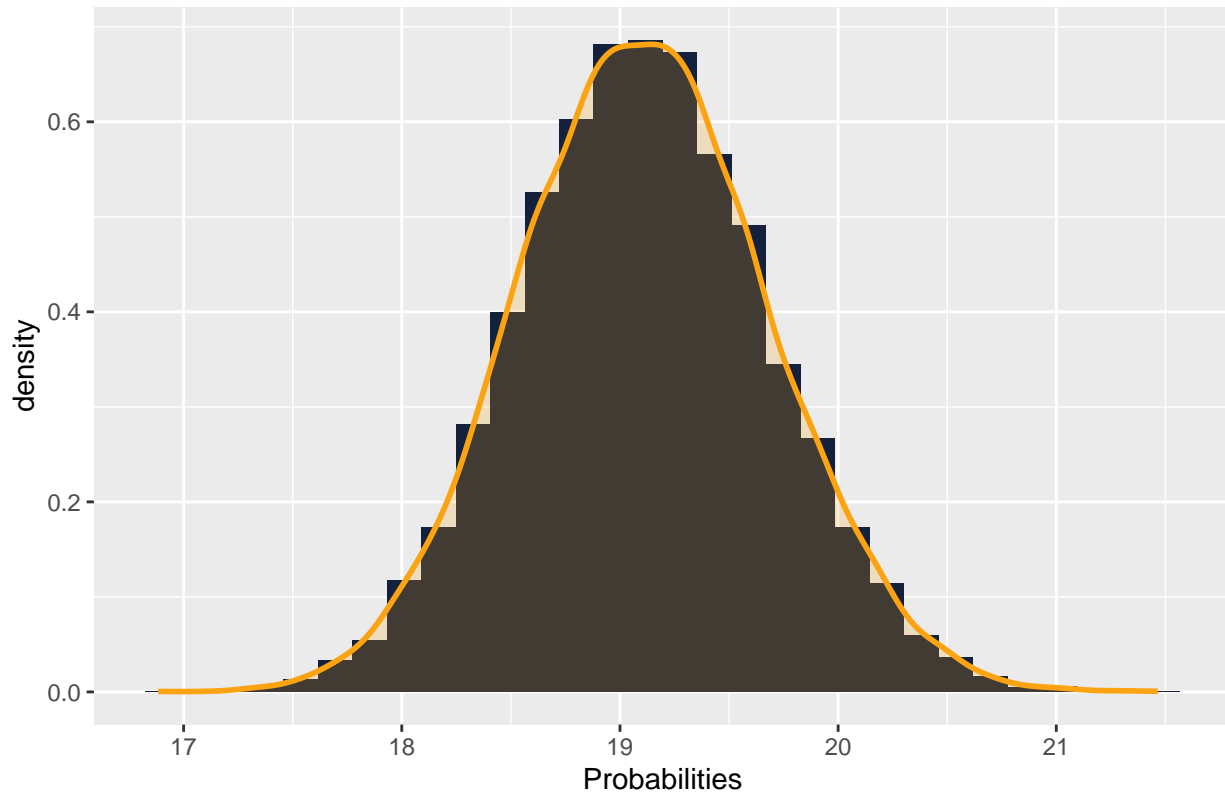
##### Infrance region A

aa=sum(reg_a$y)+40
ba=length(reg_a$y)+2
plt_df<-data.frame(x=rgamma(10000,aa,ba))

ggplot(plt_df,aes(x = x)) +
  geom_histogram(aes(y=..density..),linetype=1,fill='#14213D')+
  geom_density(alpha=.2,color="#FCA311",size=1,fill="#FCA311")+
  labs(x = 'Probabilities', y = 'density',
        title = 'Plot of the predictive distribution')
```

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```

Plot of the predictive distribution



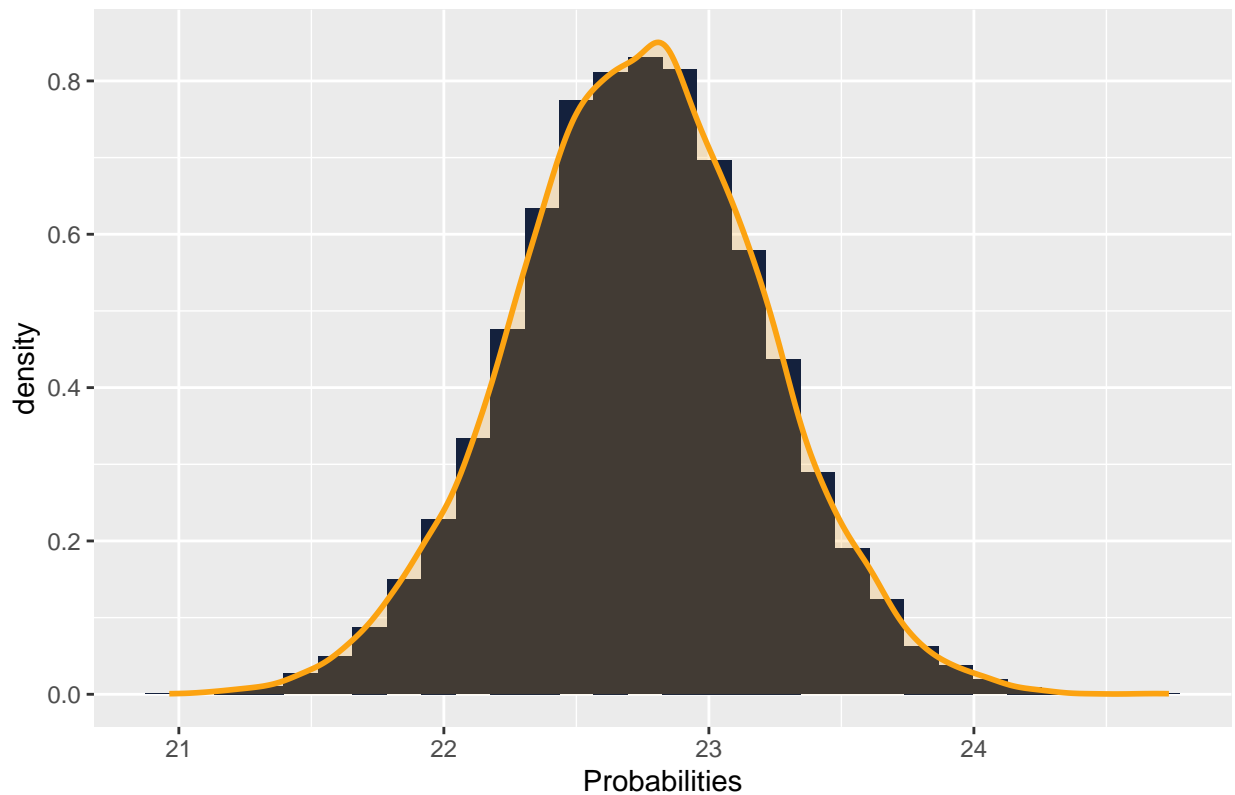
```
##### Infrance region B
```

```
ab=sum(reg_b$y)+40
bb=length(reg_b$y)+2
plt_df<-data.frame(x=rgamma(10000,ab,bb))

ggplot(plt_df,aes(x = x)) +
  geom_histogram(aes(y=..density..),linetype=1,fill='#14213D')+
  geom_density(alpha=.2,color="#FCA311",size=1,fill="#FCA311")+
  labs(x = 'Probabilities', y = 'density',
       title = 'Plot of the predictive distribution')
```

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```

Plot of the predictive distribution



```
##### future week
#a
theta_a<-rgamma(1,aa,ba)
regp_a<-c()
#b
theta_b<-rgamma(1,ab,bb)
regp_b<-c()

seq_<-seq(1,1000,1)
for (i in 1:length(seq_)) {
  regp_a[i]<-rpois(seq_[i],theta_a)
  regp_b[i]<-rpois(seq_[i],theta_b)
}
paste0("Probability that Yb>Ya is ",mean(regp_b>regp_a))
```

```
## [1] "Probability that Yb>Ya is 0.668"
```

## C

we can prove this by calculating the mean number of wolvs in both regions and then we can compare the delt mean between the 2 regions. We can see that the avg number of wolves per week in region b is higher by 18% than region A.

```
mean(regp_b)/mean(regp_a)-1
```

```
## [1] 0.1698705
```