

Statistique bayésienne avec R - exercice Monte Carlo

Julien JACQUES

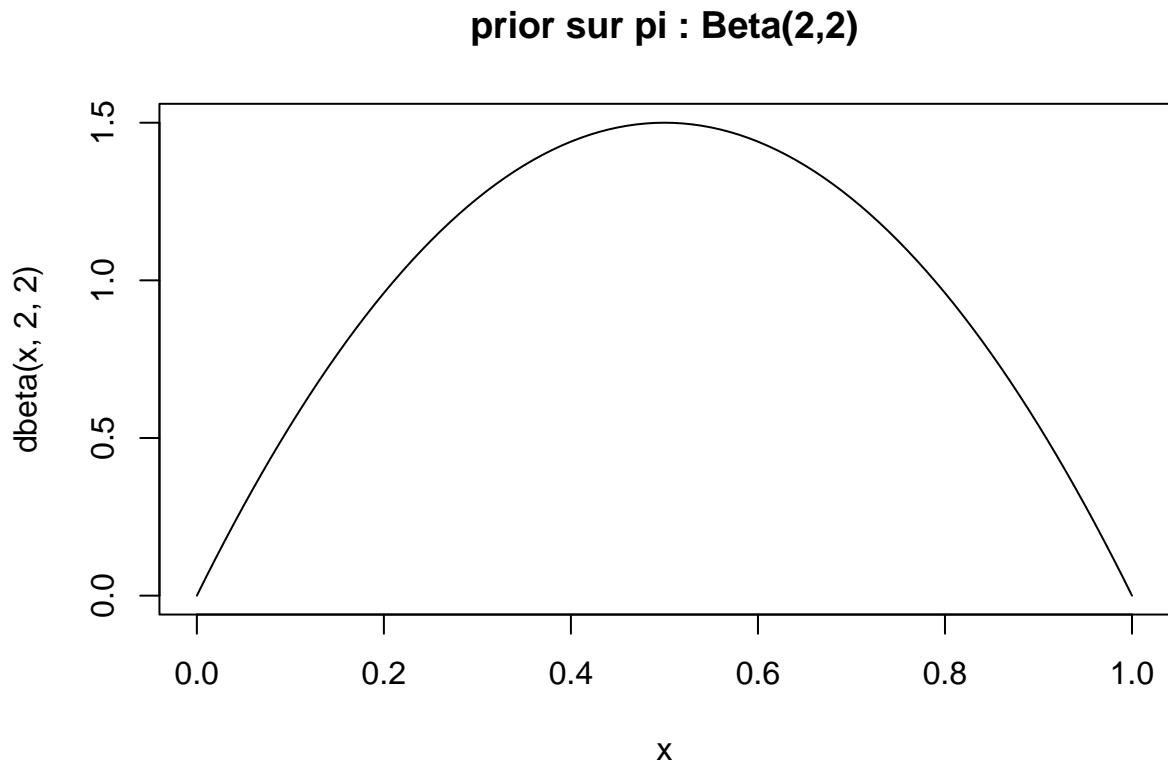
On s'intéresse au nombre Y de face obtenus sur 10 lancers d'une pièce. On suppose avoir observé 7 faces.
Considérons le modèle suivant :

$$Y|\pi \sim \mathcal{B}(10, \pi)$$

$$\pi \sim \text{Beta}(2, 2)$$

La loi a priori suppose que la pièce est non truquée.

```
x=seq(0,1,0.01)
plot(x,dbeta(x,2,2),main="prior sur pi : Beta(2,2)",type='l')
```



On obtient 7 faces.

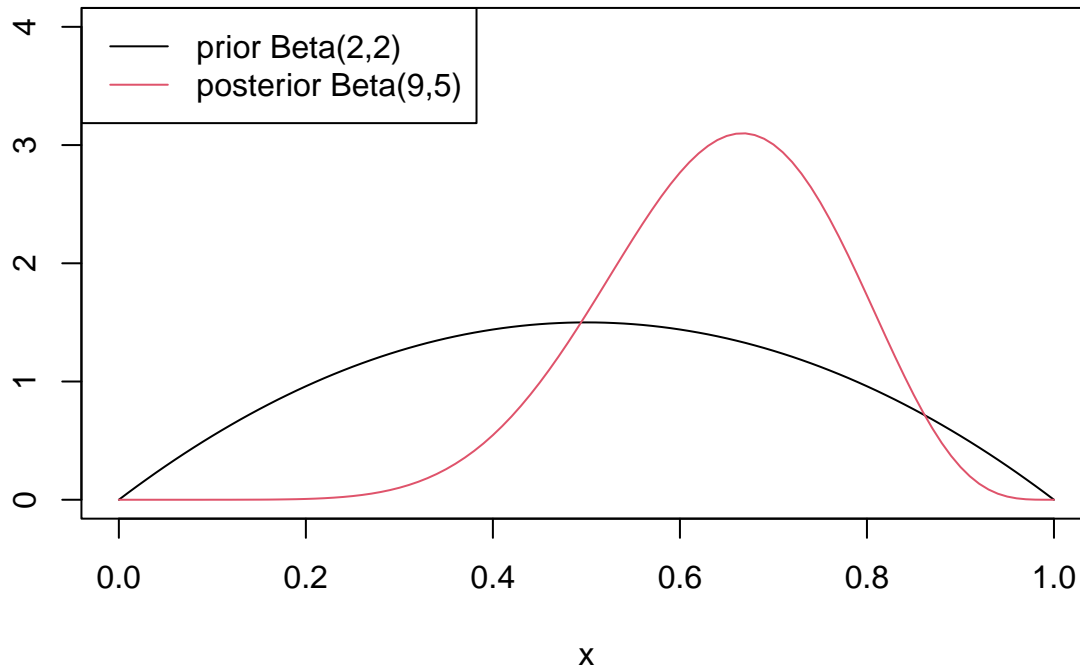
```
data=7
```

La loi a posteriori est alors :

$$Y|\pi \sim \mathcal{B}(10, \pi)$$

$$\pi|Y = 7 \sim \mathcal{B}(9, 5)$$

```
x=seq(0,1,0.01)
plot(x,dbeta(x,2,2),type='l',ylim=c(0,4),ylab="")
lines(x,dbeta(x,9,5),col=2)
legend("topleft",col=1:2,legend=c('prior Beta(2,2)', 'posterior Beta(9,5)'),lty=1)
```



Imaginons qu'on ne sache pas faire le calcul explicite de la loi a posteriori :

$$p(\pi|y) = \frac{\ell(y|\pi)p(\pi)}{\int \ell(y|\pi)p(\pi)d\pi}$$

On a :

- $\ell(\pi|y)$ est la densité d'une loi binomiale $\mathcal{B}(10, \pi)$ en y
- $p(\pi)$ est la densité d'une loi $\text{Beta}(2, 2)$ en π

Monte Carlo simple

Le numérateur est alors calculable par, pour une valeur p donnée de π :

```
dbinom(9,10,p)*dbeta(p,2,2)
```

Nous allons approcher le dénominateur par Monte-Carlo:

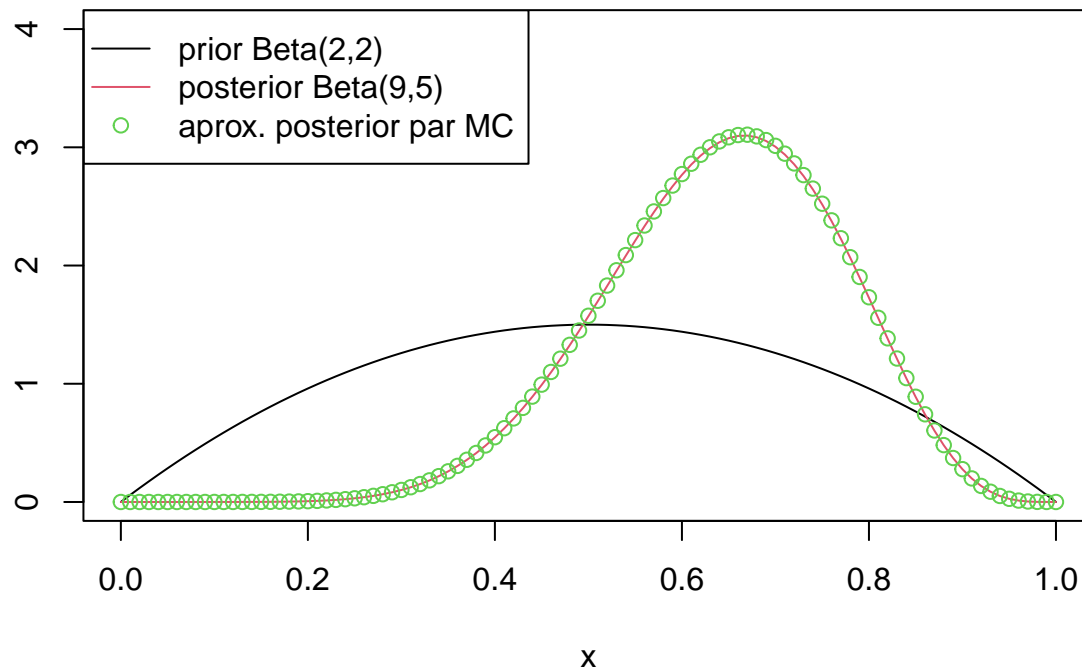
```
pi=rbeta(10000,2,2)
denom=mean(dbinom(data,10,pi))
```

Nous pouvons alors calculer la loi a posteriori :

```
post <- function(p){
  dbinom(data,10,p)*dbeta(p,2,2) / mean(dbinom(data,10,pi))
}
```

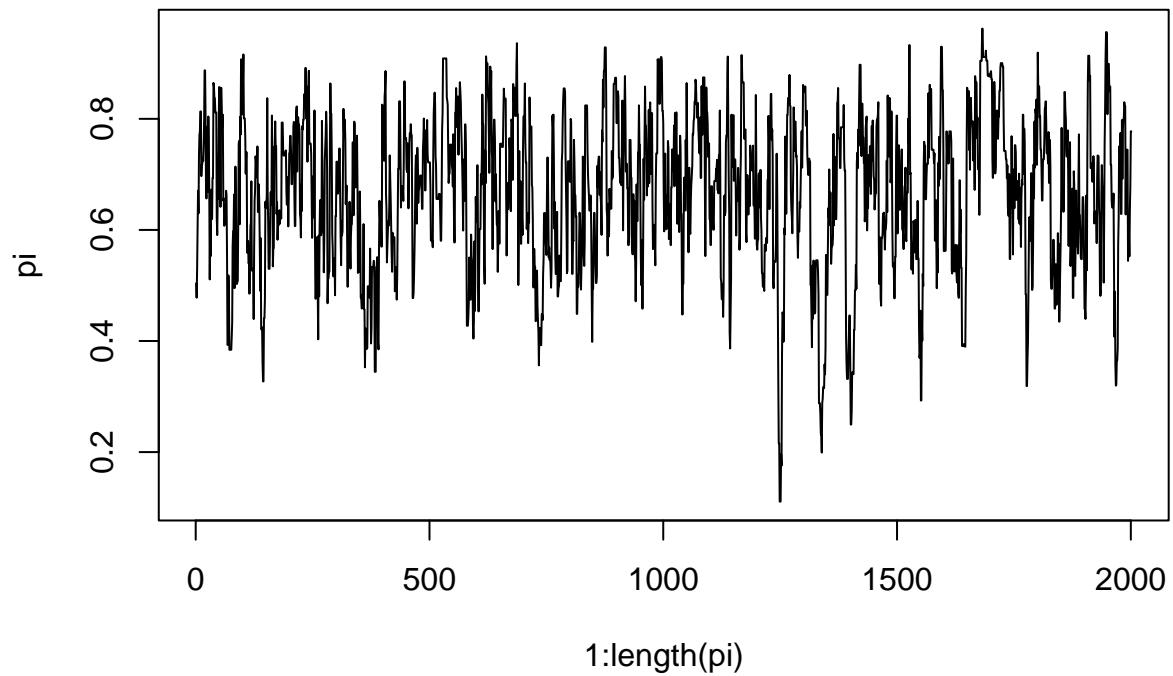
```
x=seq(0,1,0.01)
plot(x,dbeta(x,2,2),type='l',ylim=c(0,4),ylab="")
lines(x,dbeta(x,9,5),col=2)
```

```
points(x,post(x),col=3)
legend("topleft",col=1:3,legend=c('prior Beta(2,2)', 'posterior Beta(9,5)', 'aprox. posterior par MC'),lt
```



Monte Carlo par chaine de Markov

```
pi=runif(1)
sigma=.1
likeli<-function(x,theta){
  return(prod(dbinom(x,size=10,prob=theta)))
}
for (i in 1:2000){
  proposal=min(max(0,rnorm(1,pi[i],sigma)),1) #on s'assure que la proposition est dans [0,1]
  proba=min(1,likeli(data,proposal)*dnorm(pi[i],proposal,sigma)/(likeli(data,pi[i])*dnorm(proposal,pi[i]))
  tmp=rbinom(1,size=1,prob=proba)
  if (tmp){
    pi[i+1]=proposal
  }else{
    pi[i+1]=pi[i]
  }
}
plot(1:length(pi),pi,type='l')
```



```
hist(pi[1001:2000],xlim=c(0,1),freq=FALSE)  
x=seq(0,1,0.01)  
lines(x,dbeta(x,9,5),col=2)
```

Histogram of pi[1001:2000]

