Computational Statistics. Chapter 5: MCMC. Solution of exercises

Thierry Denoeux

2023-01-19

```
set.seed(2021)
```

Exercise 1

As the density of ϵ is symmetric, the MH ratio is the ratio of the densities at x^* and $x^{(t-1)}$, i.e., we have

$$R(x^{(t-1)}, x^*) = \frac{f(x^*)}{f(x^{(t-1)})} = \exp(|x^{(t-1)}| - |x^*|).$$

The following function `MH_Laplace` implements the random walk MH algorithm for this problem:

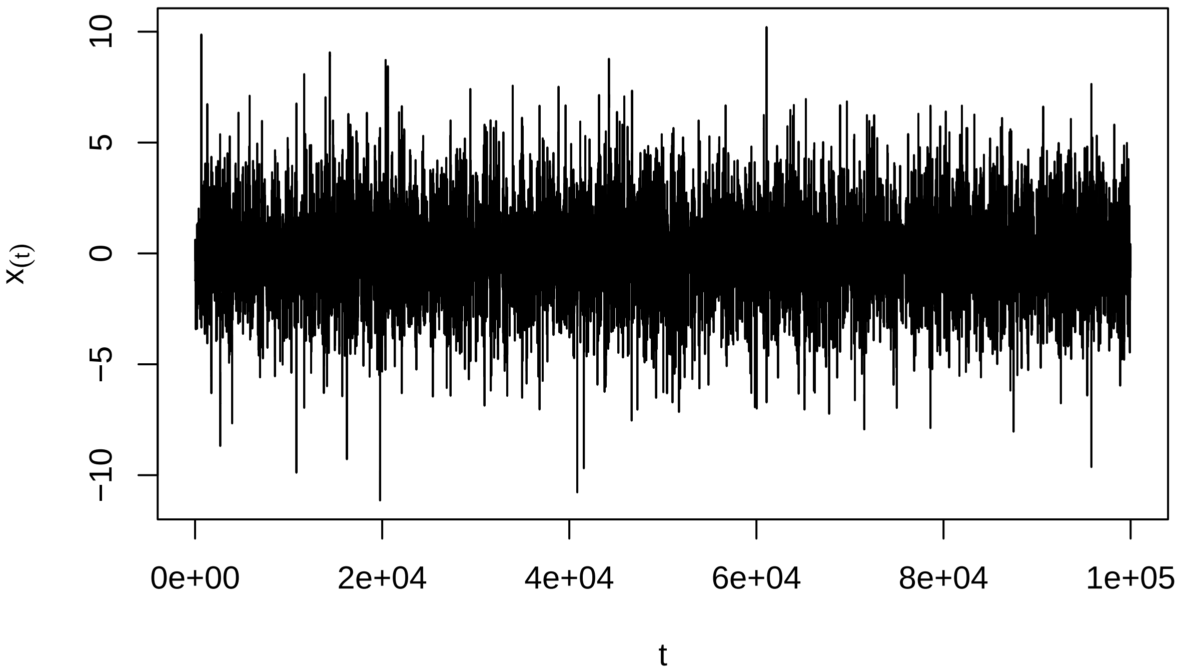
```
MH_Laplace <- function(N,sig){
  x<-vector(N,mode="numeric")
  x[1]<-rnorm(1,mean=0,sd=sig)
  for(t in (2:N)){
    epsilon<-rnorm(1,mean=0,sd=sig)
    xstar<-x[t-1]+ epsilon
    U<-runif(1)
    R<-exp(abs(x[t-1]) - abs(xstar))
    if(U <= R) x[t]<-xstar else x[t]<-x[t-1]
  }
  return(x)
}
```

Let us generate a sample of size 10^5 with $\sigma = 10$:

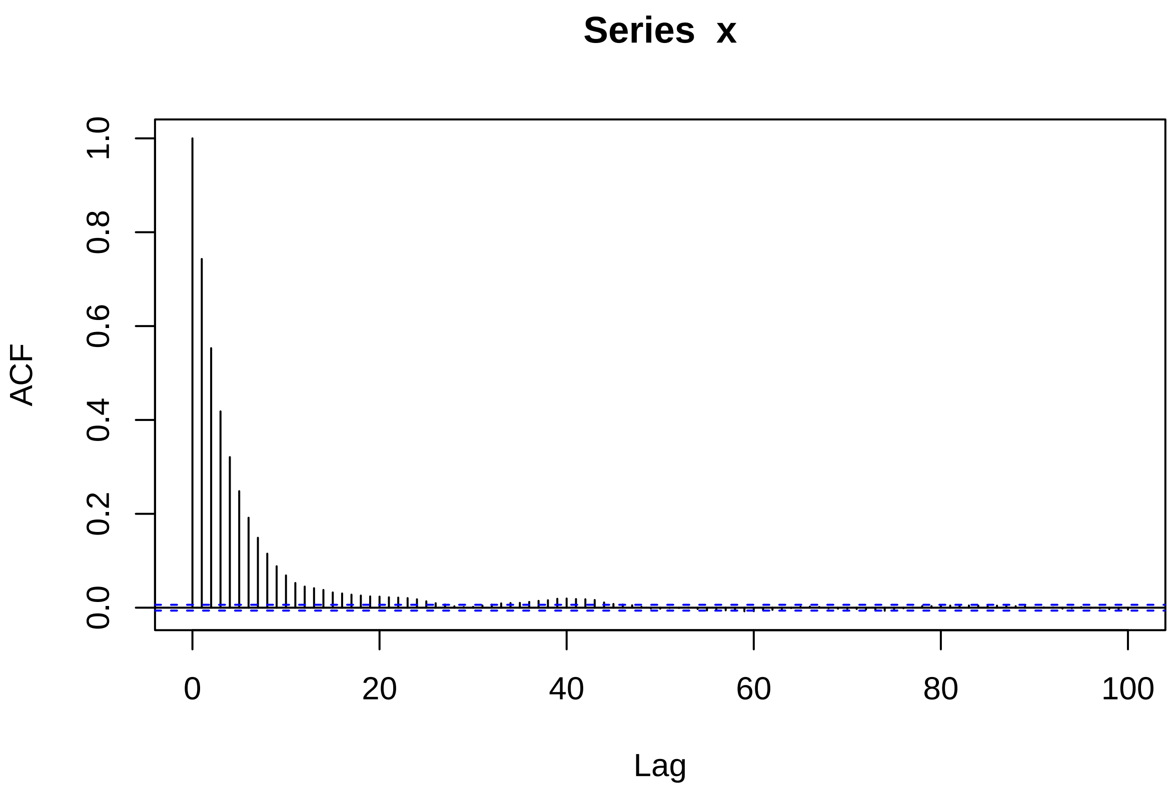
```
x<-MH_Laplace(100000,10)
```

The sample path and correlation plots show good mixing (the chain quickly moves away from its starting value, and the autocorrelation decreases quickly as the lag between iterations increases):

```
plot(x,type="l",xlab='t',ylab=expression(x[(t)]))
```

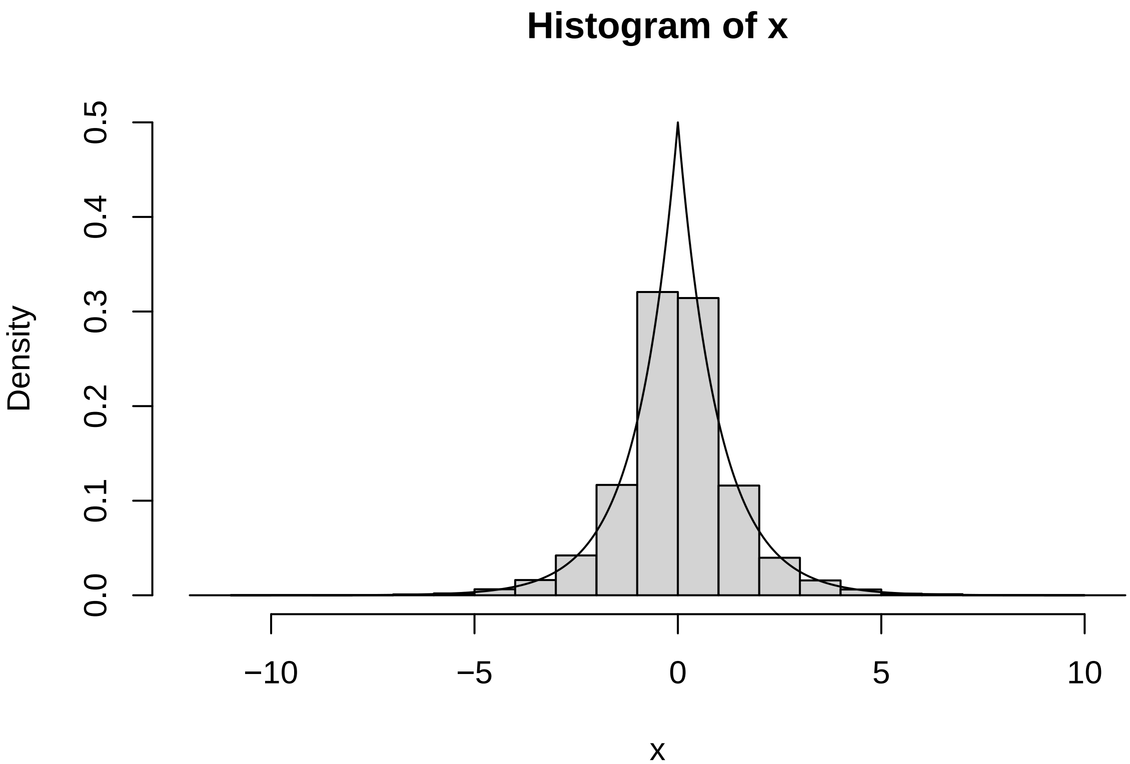



```
acf(x,lag.max=100)
```



Plot of the histogram with the Laplace density:

```
u<-seq(-10,10,0.01)
fu<-0.5*exp(-abs(u))
hist(x,freq=FALSE,ylim=range(fu))
lines(u,0.5*exp(-abs(u)))
```

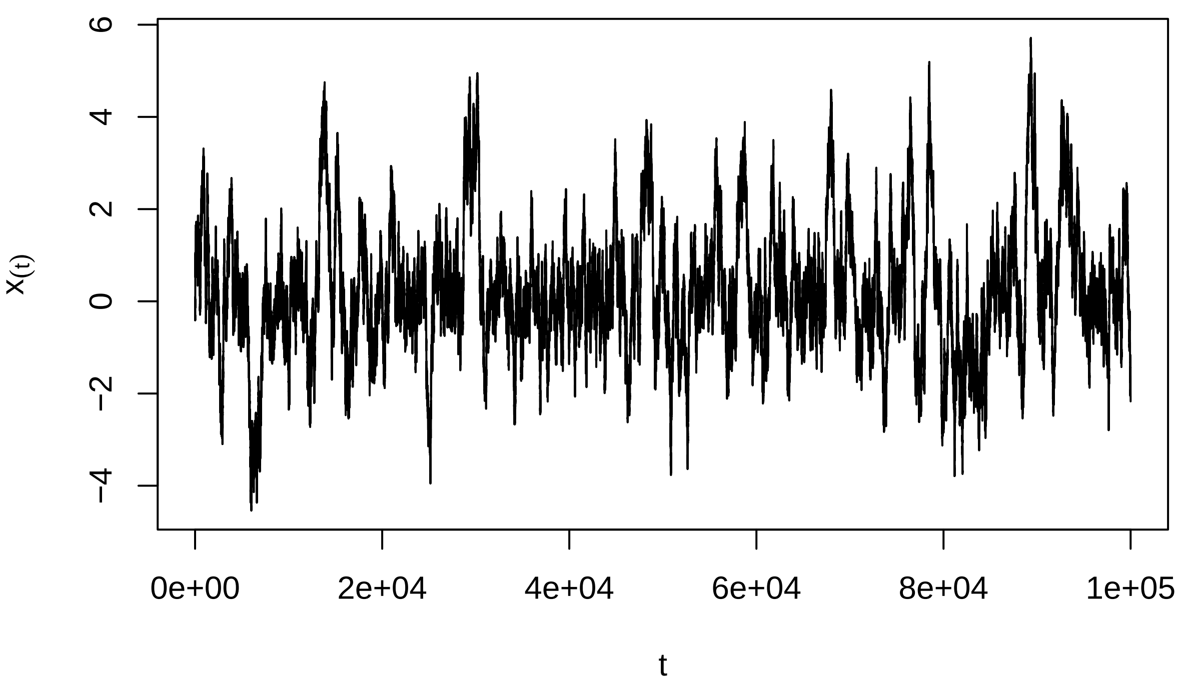


Let us now generate another sample of the same size, this time with $\sigma = 0.1$:

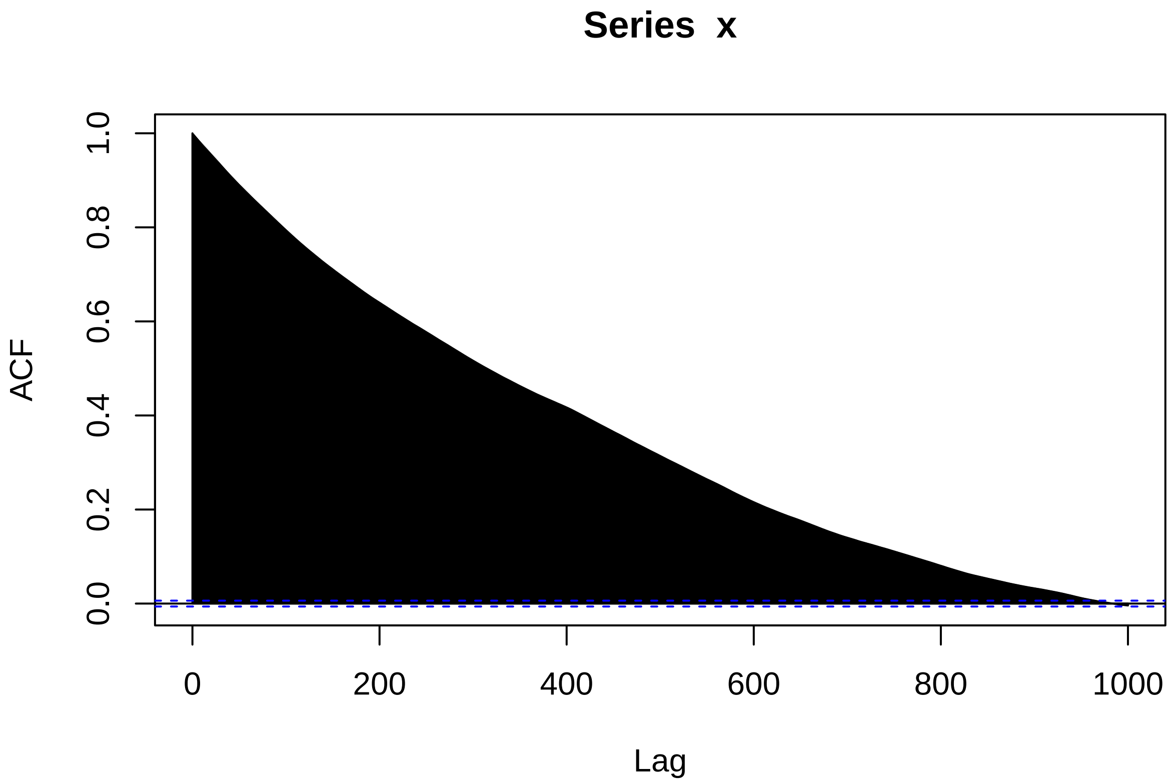
```
x<-MH_Laplace(100000,0.1)
```

This time, the sample path and correlation plots show poor mixing (the chain remains at or near the same value for many iterations, and the autocorrelation decays very slowly):

```
plot(x,type="l",xlab='t',ylab=expression(x[(t)]))
```



```r
acf(x,lag.max=1000)
```



Exercise 2

Question a

The likelihood function is

$$L(\beta; y_1, \dots, y_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(y_i - \beta x_i)^2\right) = (2\pi)^{-n/2} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}(y_i - \beta x_i)^2\right).$$

The density of the Gamma distribution with shape parameter a and rate b is $f(\beta) \propto \beta^{a-1} \exp(-b\beta) I(\beta > 0)$. Here $a = 2$ and $b = 1$, so $f(\beta) \propto \beta \exp(-\beta) I(\beta > 0)$. Consequently, the posterior density is

$$f(\beta \mid y_1, \dots, y_n) \propto \beta \exp(-\beta) \exp\left(-\frac{1}{2}\sum_{i=1}^{n}(y_i - \beta x_i)^2\right) I(\beta > 0).$$

Question b

We first write a function that computes the likelihood:

```r
loglik <- function(beta,x,y){
    n<- length(x)
    return(-0.5 * sum((y-beta*x)^2) - n/2*log(2*pi))
}
```

We then write a function that generates a MC of size N for a given data set: